```
                      GenCore version 5.1.4_p5_4578
             Copyright (c) 1993 - 2003  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         April 25, 2003, 06:57:10 ; Search time 13 Seconds
                                           (without alignments)
                                           20.370 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters:      66399

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                     %
Result              Query
  No.     Score    Match Length DB  ID                         Description
  -------------------------------------------------------------------------
     1       29     56.9      9  4  US-09-492-543-158           Sequence 158, App
     2       28     54.9      9  3  US-09-162-368B-26           Sequence 26, Appl
     3       28     54.9      9  4  US-09-161-877B-26           Sequence 26, Appl
     4       27     52.9      8  2  US-08-737-085A-6            Sequence 6, Appli
     5       27     52.9      8  3  US-09-246-258-6             Sequence 6, Appli
     6       27     52.9      8  4  US-09-532-106-6             Sequence 6, Appli
     7       27     52.9      8  4  US-09-839-666-6             Sequence 6, Appli
     8       26     51.0      6  2  US-08-672-805-12            Sequence 12, Appl
     9       26     51.0      8  4  US-08-540-922D-8            Sequence 8, Appli
    10       25     49.0      8  5  PCT-US94-03744-6            Sequence 6, Appli
    11       25     49.0      9  3  US-09-162-368B-28           Sequence 28, Appl
    12       25     49.0      9  4  US-09-161-877B-28           Sequence 28, Appl
    13       24     47.1      7  1  US-08-261-525A-6            Sequence 6, Appli
    14       24     47.1      7  4  US-09-367-940A-1            Sequence 1, Appli
    15       24     47.1      9  2  US-08-417-174-13            Sequence 13, Appl
    16       24     47.1      9  2  US-08-231-565A-13           Sequence 13, Appl
    17       24     47.1      9  2  US-09-007-961-13            Sequence 13, Appl
    18       24     47.1      9  3  US-08-159-339A-393          Sequence 393, App
    19       24     47.1      9  3  US-09-162-368B-24           Sequence 24, Appl
    20       24     47.1      9  4  US-09-161-877B-24           Sequence 24, Appl
    21       24     47.1      9  4  US-08-786-455B-5            Sequence 5, Appli
    22       24     47.1      9  4  US-09-267-439-13            Sequence 13, Appl
    23       23     45.1      7  6  5514590-12                 Patent No. 5514590
    24       23     45.1      8  2  US-08-177-109A-24           Sequence 24, Appl
    25       23     45.1      8  2  US-08-687-706-24            Sequence 24, Appl
    26       23     45.1      8  4  US-08-444-818-433           Sequence 433, App
    27       23     45.1      8  4  US-08-444-818-434           Sequence 434, App
```

```
    28       23     45.1      8  5  PCT-US94-01321-70           Sequence 70, Appl
    29       22     43.1      6  2  US-08-672-805-11            Sequence 11, Appl
    30       22     43.1      8  4  US-09-484-318-7             Sequence 7, Appli
    31       22     43.1      8  4  US-09-484-319-7             Sequence 7, Appli
    32       22     43.1      8  4  US-09-484-320-7             Sequence 7, Appli
    33       22     43.1      8  4  US-09-484-321-7             Sequence 7, Appli
    34       22     43.1      8  4  US-09-484-323-7             Sequence 7, Appli
    35       22     43.1      8  4  US-09-325-769-8             Sequence 8, Appli
    36       22     43.1      8  4  US-09-636-170-7             Sequence 7, Appli
    37       22     43.1      8  4  US-09-637-518-7             Sequence 7, Appli
    38       22     43.1      9  3  US-09-162-368B-25           Sequence 25, Appl
    39       22     43.1      9  4  US-09-161-877B-25           Sequence 25, Appl
    40       22     43.1      9  4  US-09-492-543-180           Sequence 180, App
    41       21     41.2      5  1  US-07-880-216-1             Sequence 1, Appli
    42       21     41.2      7  2  US-08-968-676-158           Sequence 158, App
    43       21     41.2      7  4  US-09-367-940A-2            Sequence 2, Appli
    44       21     41.2      7  5  PCT-US94-01321-69           Sequence 69, Appl
    45       21     41.2      8  1  US-08-594-447-37            Sequence 37, Appl


                                  ALIGNMENTS


RESULT 1
US-09-492-543-158
; Sequence 158, Application US/09492543A
; Patent No. 6316213
; GENERAL INFORMATION:
;  APPLICANT:  O'Brien, Timothy J.
;  TITLE OF INVENTION:  Compositions and Methods for the Early Diagnosis of
;  TITLE OF INVENTION:  Ovarian Cancer
;  FILE REFERENCE:  D6223CIP-B
;  CURRENT APPLICATION NUMBER: US/09/492,543A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:   09/039,211
;  PRIOR FILING DATE:   03-14-1998
;  NUMBER OF SEQ ID NOS:   189
;  SOFTWARE:   WORD 6.0.1 for Macintosh
; SEQ ID NO 158
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Residues 96-104 of the PUMP-1 protein
US-09-492-543-158

  Query Match             56.9%;  Score 29;  DB 4;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 2e+05;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        2 MFPNAP 7
            :|||:|
Db        2 LFPNSP 7


RESULT 2
US-09-162-368B-26
; Sequence 26, Application US/09162368B
; Patent No. 6083703
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;    TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;    TITLE OF INVENTION:  LYMPHOCYTES
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MICROSOFT WORD 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/162,368B
;      FILING DATE:  28-SEPT-1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/725,736
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4243US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  UNKNOWN
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-162-368B-26

  Query Match             54.9%;  Score 28;  DB 3;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 2e+05;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        2 MFPNAPY 8
            :||  ||
Db        2 LFPGRPY 8

RESULT 3
US-09-161-877B-26
; Sequence 26, Application US/09161877B
; Patent No. 6132980
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2
;    TITLE OF INVENTION:  AS A HUMAN TUMOR ANTIGEN RECOGNIZED BY
;    TITLE OF INVENTION:  CYTOTOXIC T LYMPHOCYTES
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MICROSOFT WORD 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/161,877B
;      FILING DATE:  28-SEPT-1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/725,736
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4243US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  UNKNOWN
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-161-877B-26

  Query Match             54.9%;  Score 28;  DB 4;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 2e+05;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        2 MFPNAPY 8
            :||  ||
Db        2 LFPGRPY 8

RESULT 4
US-08-737-085A-6
; Sequence 6, Application US/08737085A
; Patent No. 5869232
;  GENERAL INFORMATION:
;    APPLICANT:  SALLBERG, MATTI
;    TITLE OF INVENTION:  ANTIGEN/ANTIBODY SPECIFICITY
;    TITLE OF INVENTION:  EXCHANGER
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DARBY & DARBY PC
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/737,085A
;      FILING DATE:  27-DEC-1996
;      CLASSIFICATION:  426
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Green, Reza
;      REGISTRATION NUMBER:  38,475
;      REFERENCE/DOCKET NUMBER:  3846/0C569
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-527-7659
;      TELEFAX:  212-753-6237
;      TELEX:  236687
```

```
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-737-085A-6

  Query Match             52.9%;  Score 27;  DB 2;  Length 8;
  Best Local Similarity   83.3%;  Pred. No. 2e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 PNAPYL 9
              |||| |
Db          2 PNAPIL 7

RESULT 5
US-09-246-258-6
; Sequence 6, Application US/09246258
; Patent No. 6040137
;  GENERAL INFORMATION:
;    APPLICANT:  SALLBERG, MATTI
;    TITLE OF INVENTION:  ANTIGEN/ANTIBODY SPECIFICITY
;    TITLE OF INVENTION:  EXCHANGER
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DARBY & DARBY PC
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/246,258
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/737,085
;      FILING DATE:
;    ATTORNEY/AGENT.INFORMATION:
;      NAME:  Green, Reza
;      REGISTRATION NUMBER:  38,475
;      REFERENCE/DOCKET NUMBER:  3846/0C569
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-527-7659
;      TELEFAX:  212-753-6237
;      TELEX:  236687
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-246-258-6

  Query Match             52.9%;  Score 27;  DB 3;  Length 8;
  Best Local Similarity   .83.3%;  Pred. No. 2e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 PNAPYL 9
              |||| |
Db          2 PNAPIL 7

RESULT 6
US-09-532-106-6
; Sequence 6, Application US/09532106
; Patent No. 6245895
;   GENERAL INFORMATION:
;        APPLICANT: SALLBERG, MATTI
;        TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
;                            EXCHANGER
;        NUMBER OF SEQUENCES: 23
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DARBY & DARBY PC
;             STREET: 805 Third Avenue
;             CITY: New York
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10022
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/532,106
;             FILING DATE: 21-Mar-2000
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/737,085A
;             FILING DATE: 27-DEC-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Green, Reza
;             REGISTRATION NUMBER: 38,475
;             REFERENCE/DOCKET NUMBER: 3846/0C569
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 212-527-7659
;             TELEFAX: 212-753-6237
;             TELEX: 236687
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 8 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-532-106-6

  Query Match             52.9%;  Score 27;  DB 4;  Length 8;
  Best Local Similarity   83.3%;  Pred. No. 2e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 PNAPYL 9
              |||| |
Db          2 PNAPIL 7

RESULT 7
US-09-839-666-6
; Sequence 6, Application US/09839666
; Patent No. 6469143
;   GENERAL INFORMATION:
;        APPLICANT: SALLBERG, MATTI
;        TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
;                            EXCHANGER
;        NUMBER OF SEQUENCES: 23
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: DARBY & DARBY PC
;             STREET: 805 Third Avenue
;             CITY: New York
;             STATE: New York
;             COUNTRY: USA
;             ZIP: 10022
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
```

```
;            SOFTWARE: FastSEQ for Windows Version 2.0
;         CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/839,666
;            FILING DATE: 19-Apr-2001
;            CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 08/737,085
;            FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;            NAME: Green, Reza
;            REGISTRATION NUMBER: 38,475
;            REFERENCE/DOCKET NUMBER: 3846/0C569
;         TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 212-527-7659
;            TELEFAX: 212-753-6237
;            TELEX: 236687
;      INFORMATION FOR SEQ ID NO: 6:
;         SEQUENCE CHARACTERISTICS:
;            LENGTH: 8 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;         MOLECULE TYPE: peptide
;         SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-839-666-6

  Query Match             52.9%;  Score 27;  DB 4;  Length 8;
  Best Local Similarity   83.3%;  Pred. No. 2e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 PNAPYL 9
          |||| |
Db      2 PNAPIL 7


RESULT 8
US-08-672-805-12
; Sequence 12, Application US/08672805
; Patent No. 5831003
;  GENERAL INFORMATION:
;    APPLICANT:  Baumbach, George A.,
;    APPLICANT:  Buettner, Joseph A.,
;    APPLICANT:  Dadd, Christopher A.,
;    APPLICANT:  Hammond, David J.
;    TITLE OF INVENTION:  Peptides Which Bind to Prothrombin and
;    TITLE OF INVENTION:  Thrombin
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Bayer Corporation
;      STREET:  800 Dwight Way
;      STREET:  P. O. Box 1986
;      CITY:  Berkeley
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94701-1986
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.44Mb Storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/672,805
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Giblin, James A.
;      REGISTRATION NUMBER:  25772
;      REFERENCE/DOCKET NUMBER:  MSB-7236
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510)705-7910
```

```
;      TELEFAX:  (510)705-7904
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6
;      TYPE:  amino acid
;      STRANDEDNESS:  single strand
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  peptide
US-08-672-805-12

  Query Match             51.0%;  Score 26;  DB 2;  Length 6;
  Best Local Similarity   83.3%;  Pred. No. 2e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 PNAPYL 9
          | ||||
Db      1 PFAPYL 6


RESULT 9
US-08-540-922D-8
; Sequence 8, Application US/08540922D
; Patent No. 6284476
;  GENERAL INFORMATION:
;    APPLICANT:  Boon-Falleur, Thierry; Brichard, Vincent; Van
;    APPLICANT:  Pel, Aline; De Plaen, Etienne; Coulie, Pierre;
;    APPLICANT:  Renauld Jean-Christope; Wolfel, Thomas; and
;    APPLICANT:  Lethe, Bernard.
;    TITLE OF INVENTION:  METHOD OF IDENTIFYING INDIVIDUALS SUFFERING
;    TITLE OF INVENTION:  FROM A CELLULAR ABNORMALITY SOME OF WHOSE
;    TITLE OF INVENTION:  ABNORMAL CELLS PRESENT COMPLEXES OF HUMAN
;    TITLE OF INVENTION:  LEUKOCYTE ANTIGEN TYROSINASE DERIVED
;    TITLE OF INVENTION:  PEPTIDES, AND METHODS FOR TREATING SAID
;    TITLE OF INVENTION:  INDIVIDUALS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch 1.44 Mb storage diskette
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/540,922D
;      FILING DATE:  October 11, 1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/054,714
;      FILING DATE:  28 April 1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/994,928
;      FILING DATE:  22 December 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mary Anne Schofield
;      REGISTRATION NUMBER:  36,669
;      REFERENCE/DOCKET NUMBER:  LUD 5299.5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8
;      TYPE:  amino acids
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  Protein
```

```
;   FEATURE:
;      OTHER INFORMATION:  SEQ of aa corresponding to nt 1816-1839
;      OTHER INFORMATION:  nt of SEQ ID NO: 1
US-08-540-922D-8

  Query Match             51.0%;  Score 26;  DB 4;  Length 8;
  Best Local Similarity   57.1%;  Pred. No. 2e+05;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         3 FPNAPYL 9
             || :||:
Db         2 FPLSPYV 8


RESULT 10
PCT-US94-03744-6
; Sequence 6, Application PC/TUS9403744
;  GENERAL INFORMATION:
;    APPLICANT:  PETRI, WILLIAM A.
;    APPLICANT:  McCOY, JAMES J.
;    APPLICANT:  MANN, BARBARA J.
;    TITLE OF INVENTION:  35/31 KDA SUBUNIT OF THE
;    TITLE OF INVENTION:  ENTAMOEBA
;    TITLE OF INVENTION:  HISTOLYTICA ADHERENCE LECTIN
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  2000 Pennsylvania Avenue, Suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1812
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/03744
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/045,679
;      FILING DATE:  09-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MURASHIGE, KATE H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  29148-20005.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 887-1500
;      TELEFAX:  (202) 887-0763
;      TELEX:  90-4030
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
PCT-US94-03744-6

  Query Match             49.0%;  Score 25;  DB 5;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 2e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         4 PNAPY 8
             || ||
Db         3 PNYPY 7


RESULT 11
US-09-162-368B-28
; Sequence 28, Application US/09162368B
; Patent No. 6083703
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2 AS
;    TITLE OF INVENTION:  A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
;    TITLE OF INVENTION:  LYMPHOCYTES
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MICROSOFT WORD 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/162,368B
;      FILING DATE:  28-SEPT-1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/725,736
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4243US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  UNKNOWN
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-162-368B-28

  Query Match             49.0%;  Score 25;  DB 3;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 2e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         4 PNAPY 8
             | |||
Db         4 PGAPY 8


RESULT 12
US-09-161-877B-28
; Sequence 28, Application US/09161877B
; Patent No. 6132980
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, R.F.; ROSENBERG, S. A.
;    TITLE OF INVENTION:  IDENTIFICATION OF TRP-2
;    TITLE OF INVENTION:  AS A HUMAN TUMOR ANTIGEN RECOGNIZED BY
;    TITLE OF INVENTION:  CYTOTOXIC T LYMPHOCYTES
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
```

```
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  MICROSOFT WORD 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/161,877B
;      FILING DATE:  28-SEPT-1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/725,736
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/599,602
;      FILING DATE:  09-FEB-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KATHRYN M. BROWN
;      REGISTRATION NUMBER:  34,556
;      REFERENCE/DOCKET NUMBER:  2026-4243US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  UNKNOWN
;      TOPOLOGY:  UNKNOWN
;    MOLECULE TYPE:
;      DESCRIPTION:  PEPTIDE
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-09-161-877B-28

  Query Match             49.0%;  Score 25;  DB 4;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 2e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        4 PNAPY 8
            | |||
Db        4 PGAPY 8


RESULT 13
US-08-261-525A-6
; Sequence 6, Application US/08261525A
; Patent No. 5569598
;  GENERAL INFORMATION:
;    APPLICANT:  PARK, Soon Jae
;    APPLICANT:  LEE, Young Mee
;    APPLICANT:  WON, Teug Yeon
;    APPLICANT:  KWON, Soon Chang
;    APPLICANT:  LEE, Seung Joo
;    APPLICANT:  KIM, Jung Ho
;    APPLICANT:  KIM, Bum Joon
;    TITLE OF INVENTION:  NOVEL AMINOPEPTIDASE, PROCESSES FOR
;    TITLE OF INVENTION:  THE PREPARATION OF AMINOPEPTIDASE AND
;    TITLE OF INVENTION:  PROTEIN THEREFROM
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PARK, Soon Jae
;      STREET:  Lucky Apt. 6-101, Doryong-dong, Yuseong-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  305-340
;      ADDRESSEE:  LEE, Young Mee
;      STREET:  Shinsung Apt. 3-306, 28-4, Nae-dong, Seo-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  302-181
;      ADDRESSEE:  WON, Teug Yeon
;      STREET:  Gongjak Hanyang Apt. 3-1201, 835,
;      STREET:  Tanbang-dong, Seo-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  302-223
;      ADDRESSEE:  KWON, Soon Chang
;      STREET:  Shinsung Apt. 3-107, 28-4, Nae-dong, Seo-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  302-181
;      ADDRESSEE:  LEE, Seung Joo
;      STREET:  Lucky Apt. B-107, 386-4, Doryong-dong,
;      STREET:  Yuseong-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  305-340
;      ADDRESSEE:  KIM, Jung Ho
;      STREET:  Lucky Dormitory 511, 386-1, Doryong-dong,
;      STREET:  Yuseong-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  305-340
;      ADDRESSEE:  KIM, Bum Joon
;      STREET:  Lucky Dormitory 502, 386-1, Doryong-dong,
;      STREET:  Yuseong-gu
;      CITY:  Daejeon
;      STATE:  Daejeon
;      COUNTRY:  Republic of Korea
;      ZIP:  305-340
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk, 3.5 inch, 1.44MB storage
;      COMPUTER:  IBM PC/AT
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/261,525A
;      FILING DATE:  17-JUN-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  KR 93-11107
;      FILING DATE:  17-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shahan Islam
;      REGISTRATION NUMBER:  32,507
;      REFERENCE/DOCKET NUMBER:  A-9883
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 278-1000
;      TELEFAX:  (212) 953-7249
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-261-525A-6

  Query Match             47.1%;  Score 24;  DB 1;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 2e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        2 MFPNAP 7
```

```
            |||  |
Db      1 MFPTEP 6


RESULT 14
US-09-367-940A-1
; Sequence 1, Application US/09367940A
; Patent No. 6428997
;   GENERAL INFORMATION:
;        APPLICANT: LG CHEMICAL LTD.
;                   LEE, Young-Phil
;                   HAN, Kyuboem
;                   KIM, Se-Hoon
;                   PARK, Soon-Jae
;                   LEE, Seung-Joo
;        TITLE OF INVENTION: Aminopeptidase
;                            derived from Bacillus licheniformis
;                            and process for proparation of natural
;                            type proteins
;        NUMBER OF SEQUENCES: 4
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: BACHMAN & LAPOINTE, P.C.
;             STREET: Suite 1201, 900 Chapel Street
;             CITY: New Haven
;             STATE: Connecticut
;             COUNTRY: U.S.A.
;             ZIP: 06510-2802
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
;             COMPUTER: IBM
;             OPERATING SYSTEM: WINDOWS 95/98
;             SOFTWARE: MS WORD
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/367,940A
;             FILING DATE: 18-Aug-1999
;   INFORMATION FOR SEQ ID NO: 1 :
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 7 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1
US-09-367-940A-1

  Query Match             47.1%;  Score 24;  DB 4;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 2e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 MFPNAP 7
          |||  |
Db      1 MFPTEP 6


RESULT 15
US-08-417-174-13
; Sequence 13, Application US/08417174
; Patent No. 5844075
;  GENERAL INFORMATION:
;    APPLICANT:  KAWAKAMI, YUTAKA; ROSENBERG,
;    APPLICANT:  STEVEN A.
;    TITLE OF INVENTION:  MELANOMA ANTIGENS AND
;    TITLE OF INVENTION:  THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  126
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/417,174
;      FILING DATE:  05-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/231,565
;      FILING DATE:  22-APR-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CAROL M. GRUPPI
;      REGISTRATION NUMBER:  37,341
;      REFERENCE/DOCKET NUMBER:  2026-4124US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9
;      TYPE:  amino acid
;      STRANDEDNESS:  Unknown
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  Peptide
US-08-417-174-13

  Query Match             47.1%;  Score 24;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 PNAP 7
          ||||
Db      4 PNAP 7


Search completed: April 25, 2003, 06:59:31
Job time : 14 secs
```

```
                    GenCore version 5.1.4_p5_4578
             Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         April 25, 2003, 06:51:39 ; Search time 14 Seconds
                                           (without alignments)
                                           61.801 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters:      789

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_73:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                             SUMMARIES
                 %
Result          Query
  No.   Score   Match Length DB  ID                 Description
--------------------------------------------------------------------------
    1      19    37.3      9  2  S70345             amine oxidase (cop
    2      18    35.3      9  2  D48186             ATPase R1 subunit
    3      17    33.3      7  2  PH1408             Ig heavy chain V r
    4      17    33.3      9  2  JS0302             xenopsin-related p
    5      17    33.3      9  2  A60320             xenopsin-related p
    6      17    33.3      9  2  PL0139             carbon-monoxide de
    7      16    31.4      7  2  S21230             dermorphin (Trp-4,
    8      16    31.4      8  2  I57018             gene Cftr protein
    9      16    31.4      9  2  A44787             calliFMRFamide 10
   10      16    31.4      9  2  D41978             calliFMRFamide 4 -
   11      15    29.4      6  4  S15596             orf 3 rara 5'-regi
   12      15    29.4      9  2  S66607             quinoline 2-oxidor
   13      15    29.4      9  2  A61620             locustamyotropin I
   14      15    29.4      9  2  PH1591             Ig H chain V-D-J r
   15      14    27.5      5  2  JS0319             subesophageal gang
   16      14    27.5      8  2  S21288             lectin - potato (f
   17      14    27.5      8  2  PT0030             inulinase (EC 3.2.
   18      14    27.5      8  2  D47393             neuropeptide calla
   19      14    27.5      8  2  E47393             neuropeptide calla
   20      14    27.5      8  2  B45800             serum albumin - do
   21      14    27.5      8  4  I54017             granulocyte-colony
   22      14    27.5      9  2  S66419             tetrameric protein
   23      14    27.5      9  2  S70332             endosperm protein,
   24      14    27.5      9  2  A29477             diuretic neuropept
   25      14    27.5      9  2  S39766             cardioactive pepti
   26      14    27.5      9  2  S77984             cytochrome-c oxida
   27      14    27.5      9  2  S66635             alpha-2-macroglobu
   28      14    27.5      9  2  S10784             enamelin i - bovin
   29      13    25.5      4  2  S53508             starvation-induced
```

```
   30      13    25.5      6  2  A61049             halo-toxin - Pseud
   31      13    25.5      6  2  A44916             mosquitocidal toxi
   32      13    25.5      7  2  A15398             choline oxidase (E
   33      13    25.5      7  2  I50210             gene c-rel protein
   34      13    25.5      7  2  E48394             glycoprotein compo
   35      13    25.5      8  2  S08995             hypertrehalosemic
   36      13    25.5      8  2  S08996             hypertrehalosemic
   37      13    25.5      8  2  A49823             adipokinetic hormo
   38      13    25.5      8  2  B49823             adipokinetic hormo
   39      13    25.5      8  2  A44960             neuropeptide Led-C
   40      13    25.5      8  2  B44960             neuropeptide Led-C
   41      13    25.5      8  2  A43976             hypertrehalosemic
   42      13    25.5      8  2  B43976             hypertrehalosemic
   43      13    25.5      8  2  PH1407             Ig heavy chain V r
   44      13    25.5      8  2  PL0184             capsid protein VP-
   45      13    25.5      8  2  B39745             endoglycosylcerami

                               ALIGNMENTS


RESULT 1
S70345
amine oxidase (copper-containing) (EC 1.4.3.6) II - Aspergillus niger (fragments)
C;Species: Aspergillus niger
C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C;Accession: S70345
R;Frebort, I.; Pec, P.; Luhova, L.; Toyama, H.; Matsushita, K.; Hirota, S.; Kitagawa,
Biochim. Biophys. Acta 1295, 59-72, 1996
A;Title: Two amine oxidases from Aspergillus niger AKU 3302 contain topa quinone as t
A;Reference number: S70344; MUID:96283794; PMID:8679675
A;Accession: S70345
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-5;6-9 <FRE>
C;Keywords: oxidoreductase

  Query Match             37.3%;  Score 19;  DB 2;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 PNAPY 8
              || |
Db          5 PNXEY 9


RESULT 2
D48186
ATPase R1 subunit - wood tobacco (fragment)
C;Species: Nicotiana sylvestris (wood tobacco)
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 23-Feb-1997
C;Accession: D48186
R;De Paepe, R.; Forchioni, A.; Chetrit, P.; Vedel, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 5934-5938, 1993
A;Title: Specific mitochondrial proteins in pollen: presence of an additional ATP syn
A;Reference number: A48186; MUID:93317598; PMID:8327463
A;Accession: D48186
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <DEl>
A;Experimental source: pollen
A;Note: sequence extracted from NCBI backbone (NCBIP:134871)

  Query Match             35.3%;  Score 18;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 APY 8
              |||
Db          6 APY 8


RESULT 3
```

```
PH1408
Ig heavy chain V region - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C;Accession: PH1408; PH1405
R;Shirasawa, T.; Miyazoe, I.; Hagiwara, S.; Kimoto, H.; Shigemoto, K.; Taniguchi, M.; Ta
J. Exp. Med. 176, 1209-1214, 1992
A;Title: Heavy chain variable (VH) region diversity generated by VH gene replacement in
ia virus.
A;Reference number: PH1403; MUID:93018837; PMID:1402663
A;Accession: PH1408
A;Molecule type: DNA
A;Residues: 1-7 <SHI>
A;Experimental source: clone micro m+ 46-12-2
A;Accession: PH1405
A;Molecule type: DNA
A;Residues: 1-7 <SHI2>
A;Experimental source: clone micro m+ 46-6
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

  Query Match             33.3%;  Score 17;  DB 2;  Length 7;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RMFP 4
              | ||
Db          4 RRFP 7


RESULT 4
JS0302
xenopsin-related peptide 2 - turkey
N;Contains: xenopsin-related peptide 1
C;Species: Meleagris gallopavo (common turkey)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 18-Aug-2000
C;Accession: JS0302
R;Carraway, R.E.; Cochrane, D.E.; Mitra, S.P.
Regul. Pept. 22, 303-314, 1988
A;Title: Xenopsin-related peptide generated in avian gastric extracts.
A;Reference number: JS0302; MUID:89042995; PMID:2460902
A;Accession: JS0302
A;Molecule type: protein
A;Residues: 1-9 <CAR>
C;Comment: The peptides are present within several tissues primarily in large molecular
C;Superfamily: yeast coatomer complex alpha chain; WD repeat homology
C;Keywords: neuropeptide
F;1-9/Product: xenopsin-related peptide 2 #status experimental <XP2>
F;2-9/Product: xenopsin-related peptide 1 #status experimental <XP1>

  Query Match             33.3%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 PNAPYL 9
              |  |::
Db          3 PKRPWI 8


RESULT 5
A60320
xenopsin-related peptide 2 - rat
N;Contains: xenopsin-related peptide 1
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 18-Aug-2000
C;Accession: A60320
R;Carraway, R.E.; Mitra, S.P.; Muraki, K.
Regul. Pept. 29, 229-239, 1990
A;Title: Isolation and structures of xenopsin-related peptides from rat stomach, liver a
A;Reference number: A60320; MUID:91018491; PMID:2217904
A;Accession: A60320
A;Molecule type: protein
A;Residues: 1-9 <CAR>
A;Note: the authors purified these peptides from pepsin-treated extracts of stomach,
C;Comment: Xenopsin and xenopsin-related peptides are similar to neurotensin in seque
C;Superfamily: yeast coatomer complex alpha chain; WD repeat homology
C;Keywords: neuropeptide
F;1-9/Product: xenopsin-related peptide 2 #status experimental <MAT1>
F;2-9/Product: xenopsin-related peptide 1 #status experimental <MAT2>

  Query Match             33.3%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 PNAPYL 9
              |  |::
Db          3 PKRPWI 8


RESULT 6
PL0139
carbon-monoxide dehydrogenase (EC 1.2.99.2) large chain - Pseudomonas carboxydoflava
C;Species: Pseudomonas carboxydoflava
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 28-Apr-1993
C;Accession: PL0139
R;Kraut, M.; Hugendieck, I.; Herwig, S.; Meyer, O.
Arch. Microbiol. 152, 335-341, 1989
A;Title: Homology and distribution of CO dehydrogenase structural genes in carboxydot
A;Reference number: PL0138; MUID:90055678; PMID:2818128
A;Accession: PL0139
A;Molecule type: protein
A;Residues: 1-9 <KRA>
A;Note: 2-Met is also found
C;Comment: Cardon-monoxide dehydrogenase consists of three polypeptide chains: large,
C;Keywords: oxidoreductase

  Query Match             33.3%;  Score 17;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 NAP 7
              |||
Db          2 NAP 4


RESULT 7
S21230
dermorphin (Trp-4, Asn-7) [validated] - two-colored leaf frog (fragment)
C;Species: Phyllomedusa bicolor (two-colored leaf frog)
C;Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 18-Aug-2000
C;Accession: S21230
R;Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Kreil,
FEBS Lett. 302, 151-154, 1992
A;Title: Identification and characterization of two dermorphins from skin extracts of
A;Reference number: S21152; MUID:92339502; PMID:1633846
A;Accession: S21230
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MIG>
C;Superfamily: dermorphin precursor; dermorphin precursor amino-terminal homology

  Query Match             31.4%;  Score 16;  DB 2;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 FPN 5
              :||
Db          5 YPN 7


RESULT 8
I57018
gene Cftr protein - mouse (fragment)
C;Species: Mus sp. (mouse)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 28-Feb-1997
C;Accession: I57018
```

```
R;Dorin, J.R.; Stevenson, B.J.; Fleming, S.; Alton, E.W.; Dickinson, P.; Porteous, D.J.
Mamm. Genome 5, 465-472, 1994
A;Title: Long-term survival of the exon 10 insertional cystic fibrosis mutant mouse is a
A;Reference number: I57018; MUID:95037043; PMID:7949729
A;Accession: I57018
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-8 <RES>
A;Cross-references: GB:S74246; NID:g710482
C;Genetics:
A;Gene: Cftr

  Query Match             31.4%;  Score 16;  DB 2;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 PNAP 7
            |::|
Db        2 PDSP 5


RESULT 9
A44787
calliFMRFamide 10 - bluebottle fly (Calliphora vomitoria)
C;Species: Calliphora vomitoria
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C;Accession: A44787
R;Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe
Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992
A;Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (desi
A;Reference number: A41978; MUID:92196111; PMID:1549595
A;Accession: A44787
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <DUV>
C;Keywords: amidated carboxyl end; neuropeptide
F;9/Modified site: amidated carboxyl end (Phe) #status experimental

  Query Match             31.4%;  Score 16;  DB 2;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        4 PNAPYL 9
            ||  ::
Db        2 PNRDFM 7


RESULT 10
D41978
calliFMRFamide 4 - bluebottle fly (Calliphora vomitoria)
C;Species: Calliphora vomitoria
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C;Accession: D41978
R;Duve, H.; Johnsen, A.H.; Sewell, J.C.; Scott, A.G.; Orchard, I.; Rehfeld, J.F.; Thorpe
Proc. Natl. Acad. Sci. U.S.A. 89, 2326-2330, 1992
A;Title: Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH-2 neuropeptides (desi
A;Reference number: A41978; MUID:92196111; PMID:1549595
A;Accession: D41978
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <DUV>
C;Keywords: amidated carboxyl end; neuropeptide
F;9/Modified site: amidated carboxyl end (Phe) #status experimental

  Query Match             31.4%;  Score 16;  DB 2;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        4 PNAPYL 9
            ||  ::
Db        2 PNQDFM 7
```

```
RESULT 11
S15596
orf 3 rara 5'-region - human
C;Species: Homo sapiens (man)
C;Date: 04-Jun-1999 #sequence_revision 04-Jun-1999 #text_change 28-Jun-1999
C;Accession: S15596
R;Brand, N.J.; Petkovich, M.; Chambon, P.
Nucleic Acids Res. 18, 6799-6806, 1990
A;Title: Characterization of a functional promoter for the human retinoic acid recept
A;Reference number: S15594; MUID:91088249; PMID:2175878
A;Accession: S15596
A;Molecule type: DNA
A;Residues: 1-6 <BRA>
A;Cross-references: EMBL:X56058; NID:g35876
A;Note: this ORF from Fig. 2 is not annotated in GenBank entry HSRARA2, release 111.0
C;Comment: This sequence is not thought to be translated.
C;Genetics:
A;Gene: GDB:RARA
A;Cross-references: GDB:120337; OMIM:180240
A;Map position: 17q12-17q12

  Query Match             29.4%;  Score 15;  DB 4;  Length 6;
  Best Local Similarity   60.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        2 MFPNA 6
            | |:|
Db        1 MAPSA 5


RESULT 12
S66607
quinoline 2-oxidoreductase beta chain - Comamonas testosteroni (fragment)
C;Species: Comamonas testosteroni
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S66607
R;Schach, S.; Tshisuaka, B.; Fetzner, S.; Lingens, F.
Eur. J. Biochem. 232, 536-544, 1995
A;Title: Quinoline 2-oxidoreductase and 2-oxo-1,2-dihydroquinoline 5,6-dioxygenase fr
A;Reference number: S66606; MUID:96035889; PMID:7556204
A;Accession: S66607
A;Molecule type: protein
A;Residues: 1-9 <SCH>
A;Experimental source: strain 63

  Query Match             29.4%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        3 FPNAPY 8
            ||   |
Db        3 FPAFAY 8


RESULT 13
A61620
locustamyotropin III - migratory locust
C;Species: Locusta migratoria (migratory locust)
C;Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 11-Jul-1997
C;Accession: A61620
R;Schoofs, L.; Holman, G.M.; Hayes, T.K.; Nachman, R.J.; Kochansky, J.P.; De Loof, A.
Insect Biochem. Mol. Biol. 22, 447-452, 1992
A;Title: Isolation, identification and synthesis of locustamyotropin III and IV, two
A;Reference number: A61620
A;Accession: A61620
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <SCH>
C;Keywords: amidated carboxyl end; neuropeptide
F;9/Modified site: amidated carboxyl end (Leu) #status experimental

  Query Match             29.4%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   44.4%;  Pred. No. 2.8e+05;
```

```
  Matches    4;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |  |  | |
Db       1 RQQPFVPRL 9


RESULT 14
PH1591
Ig H chain V-D-J region (wild-type clone 142) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C;Accession: PH1591
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A;Reference number: PH1580; MUID:93301609; PMID:8315387
A;Accession: PH1591
A;Molecule type: DNA
A;Residues: 1-9 <LEV>
A;Experimental source: bone marrow pre-B lymphocyte
C;Keywords: immunoglobulin

  Query Match             29.4%;  Score 15;  DB 2;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 APY 8
           :||
Db       5 SPY 7


RESULT 15
JS0319
subesophageal ganglion pentapeptide - house cricket
C;Species: Acheta domesticus (house cricket)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000
C;Accession: JS0319
R;Wicker, C.; Wicker, C.
Comp. Biochem. Physiol. C 88, 185-187, 1987
A;Title: Isolation and structure of a peptide isolated from the suboesophageal ganglion
A;Reference number: JS0319
A;Accession: JS0319
A;Molecule type: protein
A;Residues: 1-5 <WIC>

  Query Match             27.5%;  Score 14;  DB 2;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 APY 8
           ||:
Db       3 APF 5


Search completed: April 25, 2003, 06:58:37
Job time : 16 secs
```

```
                     GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         April 25, 2003, 06:51:04 ; Search time 25 Seconds
                                           (without alignments)
                                           14.931 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters:      231

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_40:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                   %
Result           Query
  No.    Score   Match Length DB  ID                   Description
------------------------------------------------------------------------
    1       22    43.1      9  1  UPA7_HUMAN           P30093 homo sapien
    2       18    35.3      7  1  GFRP_MOUSE           P99025 mus musculu
    3       18    35.3      7  1  UN06_PINPS           P81675 pinus pinas
    4       18    35.3      8  1  B44K_PORGI           P81886 porphyromon
    5       17    33.3      8  1  PPK2_PERAM           P82692 periplaneta
    6       17    33.3      9  1  OXYF_SCYCA           P42997 scyliorhinu
    7       17    33.3      9  1  OXYT_OCTVU           P80027 octopus vul
    8       16    31.4      9  1  FAR4_CALVO           P41859 calliphora
    9       16    31.4      9  1  FARA_CALVO           P41865 calliphora
   10       16    31.4      9  1  FLA2_TREHY           P80159 treponema h
   11       16    31.4      9  1  UPA3_HUMAN           P30089 homo sapien
   12       15    29.4      8  1  ALL7_CARMA           P81809 carcinus ma
   13       15    29.4      9  1  LMT3_LOCMI           P41489 locusta mig
   14       14    27.5      5  1  SUGA_ACHDO           P19991 acheta dome
   15       14    27.5      7  1  ALL3_CARMA           P81806 carcinus ma
   16       14    27.5      7  1  ALL4_CARMA           P81807 carcinus ma
   17       14    27.5      7  1  ALL5_CARMA           P81808 carcinus ma
   18       14    27.5      7  1  FAR1_HELTI           P41871 helisoma tr
   19       14    27.5      8  1  AL15_CARMA           P81818 carcinus ma
   20       14    27.5      8  1  AL16_CARMA           P81819 carcinus ma
   21       14    27.5      8  1  ALL3_CYDPO           P82154 cydia pomon
   22       14    27.5      8  1  ALL4_CALVO           P41840 calliphora
   23       14    27.5      8  1  ALL4_CYDPO           P82155 cydia pomon
   24       14    27.5      8  1  ALL5_CALVO           P41841 calliphora
   25       14    27.5      8  1  ALL8_CARMA           P81811 carcinus ma
   26       14    27.5      8  1  ALL9_CARMA           P81812 carcinus ma
   27       14    27.5      8  1  FUSS_FUSSO           P81010 fusarium so
   28       14    27.5      9  1  AL10_CARMA           P81813 carcinus ma
   29       14    27.5      9  1  COXE_THUOB           P80975 thunnus obe
   30       14    27.5      9  1  DNF1_LOCMI           P16339 locusta mig
   31       14    27.5      9  1  NEUX_HUMAN           P04277 homo sapien
   32       14    27.5      9  1  TKC1_CALVO           P41517 calliphora
   33       13    25.5      5  1  PAP2_PARMA           P81864 pardachirus
   34       13    25.5      7  1  CHOX_ALCSP           P16101 alcaligenes
   35       13    25.5      7  1  FAR4_PANRE           P41875 panagrellus
   36       13    25.5      7  1  MNP1_LEPDE           P42984 leptinotars
   37       13    25.5      7  1  UF04_MOUSE           P38642 mus musculu
   38       13    25.5      8  1  FAR1_PANRE           P41872 panagrellus
   39       13    25.5      8  1  HTF1_PERAM           P04548 periplaneta
   40       13    25.5      8  1  HTF2_PERAM           P04549 periplaneta
   41       13    25.5      8  1  HTF_TENMO            P25419 tenebrio mo
   42       13    25.5      8  1  PPK3_PERAM           P82618 periplaneta
   43       13    25.5      8  1  UF06_MOUSE           P38644 mus musculu
   44       13    25.5      8  1  UPA1_HUMAN           P30087 homo sapien
   45       13    25.5      9  1  CCAP_CARMA           P38556 carcinus ma

                                ALIGNMENTS

RESULT 1
UPA7_HUMAN
ID   UPA7_HUMAN     STANDARD;      PRT;     9 AA.
AC   P30093;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Unknown protein from 2D-page of plasma (Spot 18) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=93092937; PubMed=1459097;
RA   Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA   Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA   Hochstrasser D.F.;
RT   "Plasma protein map: an update by microsequencing.";
RL   Electrophoresis 13:707-714(1992).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 5.05, ITS MW IS: 37 kDa.
DR   SWISS-2DPAGE; P30093; HUMAN.
FT   NON_TER       1      1
FT   UNSURE        5      5
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1042 MW;  5C14477AEB0772C7 CRC64;

  Query Match             43.1%;  Score 22;  DB 1;  Length 9;
  Best Local Similarity   42.9%;  Pred. No. 1.1e+05;
  Matches    3; Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        2 MFPNAPY 8
            : |  ||
Db        2 LVPEXPY 8


RESULT 2
GFRP_MOUSE
ID   GFRP_MOUSE     STANDARD;      PRT;     7 AA.
AC   P99025;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).
GN   GCHFR OR GFRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Liver;
RA   Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
RA   Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
```

```
RA   Cowthorne M.;
RL   Submitted (AUG-1998) to the SWISS-PROT data bank.
CC   -!- FUNCTION: MEDIATES TETRAHYDROBIOPTERIN INHIBITION OF GTP
CC       CYCLOHYDROLASE I. THIS INHIBITION IS REVERSED BY L-PHENYLALANINE
CC       (BY SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
DR   SWISS-2DPAGE; P99025; MOUSE.
FT   INIT_MET      0      0
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  806 MW;  71B5B057273B4700 CRC64;

  Query Match             35.3%;  Score 18;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PYL 9
              |||
Db          1 PYL 3


RESULT 3
UN06_PINPS
ID   UN06_PINPS     STANDARD;      PRT;     7 AA.
AC   P81675;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Unknown protein from 2D-page of needles (N141) (Fragment).
OS   Pinus pinaster (Maritime pine).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX   NCBI_TaxID=71647;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Needle;
RX   MEDLINE=99274088; PubMed=10344291;
RA   Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA   Frigerio J.-M., Plomion C.;
RT   "Separation and characterization of needle and xylem maritime pine
RT   proteins.";
RL   Electrophoresis 20:1098-1108(1999).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 6.6, ITS MW IS: 25 kDa.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  823 MW;  69D76724486B5740 CRC64;

  Query Match             35.3%;  Score 18;  DB 1;  Length 7;
  Best Local Similarity   28.6%;  Pred. No. 1.1e+05;
  Matches    2;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 MFPNAPY 8
              :: | |:
Db          1 LYGNLPF 7


RESULT 4
B44K_PORGI
ID   B44K_PORGI     STANDARD;      PRT;     8 AA.
AC   P81886;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   44 kDa immunogenic protein (Fragment).
OS   Porphyromonas gingivalis (Bacteroides gingivalis).
OC   Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;
OC   Porphyromonadaceae; Porphyromonas.
OX   NCBI_TaxID=837;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=VPB 3492;
RX   MEDLINE=20198497; PubMed=10731616;
RA   Norris J.M., Love D.N.;
```

```
RT   "Serum antibody responses of cats to soluble whole cell antigens of
RT   feline Porphyromonas gingivalis.";
RL   Vet. Microbiol. 73:37-49(2000).
CC   -!- SIMILARITY: TO P.GINGIVALIS HEMAGGLUTININ A.
KW   Antigen.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  989 MW;  9554540326CB476D CRC64;

  Query Match             35.3%;  Score 18;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 APY 8
              |||
Db          1 APY 3


RESULT 5
PPK2_PERAM
ID   PPK2_PERAM     STANDARD;      PRT;     8 AA.
AC   P82692;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Pyrokinin-2 (Pea-PK-2) (FXPRL-amide).
OS   Periplaneta americana (American cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blattoidea; Blattidae; Periplaneta.
OX   NCBI_TaxID=6978;
RN   [1]
RP   SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC   TISSUE=Corpora cardiaca;
RX   MEDLINE=97353923; PubMed=9210163;
RA   Predel R., Kellner R., Kaufmann R., Penzlin H., Gaede G.;
RT   "Isolation and structural elucidation of two pyrokinins from the
RT   retrocerebral complex of the American cockroach.";
RL   Peptides 18:473-478(1997).
RN   [2]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=20189894; PubMed=10723010;
RA   Predel R., Eckert M.;
RT   "Tagma-specific distribution of FXPRLamides in the nervous system of
RT   the American cockroach.";
RL   J. Comp. Neurol. 419:352-363(2000).
CC   -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- TISSUE SPECIFICITY: CORPORA CARDIACA.
CC   -!- MASS SPECTROMETRY: MW=883; METHOD=MALDI.
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR   InterPro; IPR001484; Pyrokinin.
DR   PROSITE; PS00539; PYROKININ; FALSE_NEG.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  884 MW;  C834176DD9D77775 CRC64;

  Query Match             33.3%;  Score 17;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+05;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 PNAPYL 9
              | || |
Db          3 PFAPRL 8


RESULT 6
OXYF_SCYCA
ID   OXYF_SCYCA     STANDARD;      PRT;     9 AA.
AC   P42997;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Phasvatocin.
```

```
OS   Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC   Scyliorhinidae; Scyliorhinus.
OX   NCBI_TaxID=7830;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pituitary;
RX   MEDLINE=95062247; PubMed=7972045;
RA   Chauvet J., Rouille Y., Chauveau C., Chauvet M.-T., Acher R.;
RT   "Special evolution of neurohypophysial hormones in cartilaginous
RT   fishes: asvatocin and phasvatocin, two oxytocin-like peptides
RT   isolated from the spotted dogfish (Scyliorhinus caniculus).";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:11266-11270(1994).
CC   -!- FUNCTION: DISPLAYS OXYTOCIC ACTIVITY ON RAT UTERUS.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   InterPro; IPR000981; Neurhyp_horm.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1016 MW;  17EDD76EB44449DB CRC64;

  Query Match             33.3%;  Score 17;  DB 1;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 1.1e+05;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 FPNAP 7
              | | |
Db          3 FNNCP 7


RESULT 7
OXYT_OCTVU
ID   OXYT_OCTVU     STANDARD;      PRT;     9 AA.
AC   P80027;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Cephalotocin.
OS   Octopus vulgaris (Octopus).
OC   Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC   Incirrata; Octopodidae; Octopus.
OX   NCBI_TaxID=6645;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Nerve endings;
RX   MEDLINE=92270139; PubMed=1589145;
RA   Reich G.;
RT   "A new peptide of the oxytocin/vasopressin family isolated from
RT   nerves of the cephalopod Octopus vulgaris.";
RL   Neurosci. Lett. 134:191-194(1992).
CC   -!- FUNCTION: HAS A ROLE IN THE NEUROSECRETORY SYSTEM OF THE VENA
CC       CAVA.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   InterPro; IPR000981; Neurhyp_horm.
DR   InterPro; IPR001230; Prenyl_site.
DR   Pfam; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1072 MW;  17FF476EB45409DB CRC64;

  Query Match             33.3%;  Score 17;  DB 1;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 1.1e+05;
  Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 FPNAP 7
              | | |
Db          3 FRNCP 7
```

```
RESULT 8
FAR4_CALVO
ID   FAR4_CALVO     STANDARD;      PRT;     9 AA.
AC   P41859;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   CalliFMRFamide 4.
OS   Calliphora vomitoria (Blue blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Oestroidea; Calliphoridae; Calliphora.
OX   NCBI_TaxID=27454;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Thoracic ganglion;
RX   MEDLINE=92196111; PubMed=1549595;
RA   Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA   Rehfeld J.F., Thorpe A.;
RT   "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT   neuropeptides (designated calliFMRFamides) from the blowfly
RT   Calliphora vomitoria.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   PIR; D41978; D41978.
KW   Neuropeptide; Amidation.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1182 MW;  31730699CAB6D457 CRC64;

  Query Match             31.4%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 1.1e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 PNAPYL 9
              ||  ::
Db          2 PNQDFM 7


RESULT 9
FARA_CALVO
ID   FARA_CALVO     STANDARD;      PRT;     9 AA.
AC   P41865;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   CalliFMRFamide 10.
OS   Calliphora vomitoria (Blue blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC   Muscomorpha; Oestroidea; Calliphoridae; Calliphora.
OX   NCBI_TaxID=27454;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Thoracic ganglion;
RX   MEDLINE=92196111; PubMed=1549595;
RA   Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA   Rehfeld J.F., Thorpe A.;
RT   "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT   neuropeptides (designated calliFMRFamides) from the blowfly
RT   Calliphora vomitoria.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
DR   PIR; A44787; A44787.
KW   Neuropeptide; Amidation.
FT   MOD_RES       9      9       AMIDATION.
FT   UNSURE        1      1       OR S OR A.
SQ   SEQUENCE   9 AA;  1183 MW;  29D00699CAB40457 CRC64;

  Query Match             31.4%;  Score 16;  DB 1;  Length 9;
```

```
  Best Local Similarity   33.3%;  Pred. No. 1.1e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 PNAPYL 9
              ||  ::
Db          2 PNRDFM 7


RESULT 10
FLA2_TREHY
ID   FLA2_TREHY     STANDARD;      PRT;     9 AA.
AC   P80159;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Flagellar filament outer layer protein flaA2 (35 kDa sheath protein)
DE   (Fragment).
GN   FLAA2.
OS   Treponema hyodysenteriae (Serpulina hyodysenteriae).
OC   Bacteria; Spirochaetales; Brachyspiraceae; Brachyspira.
OX   NCBI_TaxID=159;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=C5;
RX   MEDLINE=93139764; PubMed=1487733;
RA   Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,
RA   van der Zeijst B.A.M., Kusters J.G.;
RT   "The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are
RT   composed of two sheath proteins and three core proteins.";
RL   J. Gen. Microbiol. 138:2697-2706(1992).
CC   -!- FUNCTION: COMPONENT OF THE OUTER LAYER OF THE FLAGELLA.
CC   -!- SUBUNIT: THE FLAGELLUM CONSISTS OF AN OUTER LAYER COMPOSED OF TWO
CC       SHEATH PROTEINS, FLAA1 (44 kDa) AND FLAA2 (35 kDa) AROUND A CORE
CC       THAT CONTAINS THREE PROTEINS FLAB1 (37 kDa), FLAB2 (34 kDa) AND
CC       FLAB3 (32 kDa).
CC   -!- SUBCELLULAR LOCATION: Periplasmic flagellum.
KW   Flagella; Periplasmic.
FT   UNSURE        2      2
FT   UNSURE        8      9
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1129 MW;  855A19C68B4772D1 CRC64;

  Query Match             31.4%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 PYL 9
              ||:
Db          4 PYM 6


RESULT 11
UPA3_HUMAN
ID   UPA3_HUMAN     STANDARD;      PRT;     9 AA.
AC   P30089;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Unknown protein from 2D-page of plasma (Spot 11) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=93092937; PubMed=1459097;
RA   Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA   Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA   Hochstrasser D.F.;
RT   "Plasma protein map: an update by microsequencing.";
RL   Electrophoresis 13:707-714(1992).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
```

```
CC       PROTEIN IS: 4.6, ITS MW IS: 46 kDa.
DR   SWISS-2DPAGE; P30089; HUMAN.
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1056 MW;  26F2B1BAF769C737 CRC64;

  Query Match             31.4%;  Score 16;  DB 1;  Length 9;
  Best Local Similarity   28.6%;  Pred. No. 1.1e+05;
  Matches    2;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 MFPNAPY 8
              :||    :
Db          3 LFPXTDF 9


RESULT 12
ALL7_CARMA
ID   ALL7_CARMA     STANDARD;      PRT;     8 AA.
AC   P81809; P81810; P81804;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Carcinustatin 7 [Contains: Carcinustatin 6; Carcinustatin 1].
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC   Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC   Brachyura; Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   CHAIN         1      8       CARCINUSTATIN 7.
FT   CHAIN         2      8       CARCINUSTATIN 6.
FT   CHAIN         4      8       CARCINUSTATIN 1.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  825 MW;  922879CDCB4775BD CRC64;

  Query Match             29.4%;  Score 15;  DB 1;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 APY 8
              :||
Db          2 SPY 4


RESULT 13
LMT3_LOCMI
ID   LMT3_LOCMI     STANDARD;      PRT;     9 AA.
AC   P41489;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Locustamyotropin 3 (LOM-MT-3).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Oedipodinae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Brain;
RA   Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., Kochansky J.P.,
RA   de Loof A.;
```

```
RT   "Isolation, identification and synthesis of locustamyotropin III and
RT   IV, two additional neuropeptides of Locusta migratoria: members of the
RT   locustamyotropin peptide family.";
RL   Insect Biochem. Mol. Biol. 22:447-452(1992).
CC   -!- FUNCTION: POTENT MEDIATOR OF VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR   InterPro; IPR001484; Pyrokinin.
DR   PROSITE; PS00539; PYROKININ; 1.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  1140 MW;  D5AE1772C9D776C6 CRC64;

  Query Match             29.4%;  Score 15;  DB 1;  Length 9;
  Best Local Similarity   44.4%;  Pred. No. 1.1e+05;
  Matches    4;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |  |  | |
Db        1 RQQPFVPRL 9


RESULT 14
SUGA_ACHDO
ID   SUGA_ACHDO     STANDARD;      PRT;     5 AA.
AC   P19991;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Suboesophageal ganglion pentapeptide.
OS   Acheta domesticus (House cricket).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC   Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Ensifera;
OC   Gryllidae; Gryllinae; Acheta.
OX   NCBI_TaxID=6997;
RN   [1]
RP   SEQUENCE.
RA   Wicker C., Wicker C.;
RT   "Isolation and structure of a peptide isolated from the
RT   suboesophageal ganglion of Acheta domesticus (orthoptera).";
RL   Comp. Biochem. Physiol. 88C:185-187(1987).
CC   -!- SUBCELLULAR LOCATION: MAIN PEPTIDE FROM THE SUBOESOPHAGEAL
CC       GANGLIA.
DR   PIR; JS0319; JS0319.
SQ   SEQUENCE   5 AA;  476 MW;  69D76DDDDDB00000 CRC64;

  Query Match             27.5%;  Score 14;  DB 1;  Length 5;
  Best Local Similarity   66.7%;  Pred. No. 1.1e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        6 APY 8
            ||:
Db        3 APF 5


RESULT 15
ALL3_CARMA
ID   ALL3_CARMA     STANDARD;      PRT;     7 AA.
AC   P81806;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Carcinustatin 3.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC   Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC   Brachyura; Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Multigene family.
SQ   SEQUENCE   7 AA;  796 MW;  672879CDCB476B70 CRC64;

  Query Match             27.5%;  Score 14;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        7 PY 8
            ||
Db        2 PY 3


Search completed: April 25, 2003, 06:58:16
Job time : 27 secs
```

```
                         GenCore version 5.1.4_p5_4578
               Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         April 25, 2003, 06:52:35 ; Search time 28 Seconds
                                           (without alignments)
                                           66.229 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters:      686

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_21:*
                 1:  sp_archea:*
                 2:  sp_bacteria:*
                 3:  sp_fungi:*
                 4:  sp_human:*
                 5:  sp_invertebrate:*
                 6:  sp_mammal:*
                 7:  sp_mhc:*
                 8:  sp_organelle:*
                 9:  sp_phage:*
                 10:  sp_plant:*
                 11:  sp_rodent:*
                 12:  sp_virus:*
                 13:  sp_vertebrate:*
                 14:  sp_unclassified:*
                 15:  sp_rvirus:*
                 16:  sp_bacteriap:*
                 17:  sp_archeap:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                   SUMMARIES
                   %
Result           Query
   No.   Score   Match Length DB  ID                      Description
-----------------------------------------------------------------------------
    1      20    39.2      9   4  Q9UMF3                  Q9umf3 homo sapien
    2      17    33.3      8   4  Q9UJ50                  Q9uj50 homo sapien
    3      17    33.3      9   4  Q9UM87                  Q9um87 homo sapien
    4      17    33.3      9   4  Q96T78                  Q96t78 homo sapien
    5      17    33.3      9  10   P82429                  P82429 nicotiana t
    6      16    31.4      8   2  Q9R7T2                  Q9r7t2 escherichia
    7      16    31.4      9   3  Q9P8E5                  Q9p8e5 kluyveromyc
    8      15    29.4      8   5  Q9TWH6                  Q9twh6 perinereis
    9      15    29.4      9   4  Q9H522                  Q9h522 homo sapien
   10      15    29.4      9   4  Q9UCS8                  Q9ucs8 homo sapien
   11      15    29.4      9  10   Q9S8J8                  Q9s8j8 oryza sativ
   12      14    27.5      7   8  O98866                  O98866 spinacia ol
   13      14    27.5      7  15   Q07624                  Q07624 rous sarcom
   14      14    27.5      8   3  P82858                  P82858 puccinia re
   15      14    27.5      8   8  Q8WFR5                  Q8wfr5 diadema pau
   16      14    27.5      8   8  Q8W8G5                  Q8w8g5 diadema ant
   17      14    27.5      8   8  Q8W8G4                  Q8w8g4 diadema mex
   18      14    27.5      8   8  Q8W8G3                  Q8w8g3 diadema pau
   19      14    27.5      8   8  Q8W8G2                  Q8w8g2 diadema sav
   20      14    27.5      8  11   Q99P40                  Q99p40 mus musculu
   21      14    27.5      9   2  P83157                  P83157 anabaena sp
   22      14    27.5      9   6  Q9TRU7                  Q9tru7 bos taurus
   23      14    27.5      9   8  Q8WFS4                  Q8wfs4 diadema mex
   24      14    27.5      9   8  Q8W8X4                  Q8w8x4 diadema mex
   25      14    27.5      9   8  Q8W8W6                  Q8w8w6 diadema ant
   26      14    27.5      9   8  Q8W8W5                  Q8w8w5 diadema set
   27      14    27.5      9   8  Q9T2K9                  Q9t2k9 spinacia ol
   28      14    27.5      9  12   Q67605                  Q67605 squash leaf
   29      14    27.5      9  12   Q67606                  Q67606 squash leaf
   30      13    25.5      8   2  Q9AGP4                  Q9agp4 arthrobacte
   31      13    25.5      8   2  Q49534                  Q49534 mycoplasma
   32      13    25.5      8   5  P83277                  P83277 macrobrachi
   33      13    25.5      9   2  Q99193                  Q99193 pseudomonas
   34      13    25.5      9   4  Q9UKJ6                  Q9ukj6 homo sapien
   35      13    25.5      9   4  Q9UC36                  Q9uc36 homo sapien
   36      13    25.5      9   6  Q28121                  Q28121 bos taurus
   37      13    25.5      9   6  P82926                  P82926 bos taurus
   38      13    25.5      9  12   O71066                  O71066 canine dist
   39      13    25.5      9  16   Q935G1                  Q935g1 salmonella
   40      12    23.5      8   4  Q9P0K3                  Q9p0k3 homo sapien
   41      12    23.5      8  10   P82324                  P82324 pisum sativ
   42      12    23.5      8  11   Q9JLD7                  Q9jld7 mesocricetu
   43      12    23.5      8  11   P82598                  P82598 rattus norv
   44      12    23.5      8  12   O90345                  O90345 hepatitis g
   45      12    23.5      8  12   Q9WJ33                  Q9wj33 pseudorabie

                                  ALIGNMENTS

RESULT 1
Q9UMF3
ID   Q9UMF3      PRELIMINARY;      PRT;     9 AA.
AC   Q9UMF3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PD-1 protein (Fragment).
GN   PD-1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=97473511; PubMed=9332365;
RA   Finger L.R., Pu J., Wasserman R., Vibhakar R., Louie E., Hardy R.R.,
RA   Burrows P.D., Billips L.G.;
RT   "The human PD-1 gene: complete cDNA, genomic organization, and
RT   developmentally regulated expression in B cell progenitors.";
RL   Gene 197:177-187(1997).
DR   EMBL; U64864; AAC51774.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1067 MW;  DD4A676DC6C76046 CRC64;

  Query Match             39.2%;  Score 20;  DB 4;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 6.7e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        4 PNAPY 8
            | ||:
Db        4 PQAPW 8


RESULT 2
Q9UJ50
ID   Q9UJ50      PRELIMINARY;      PRT;     8 AA.
AC   Q9UJ50;
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Latrophilin-2 (Fragment).
GN   LPHH1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99153747; PubMed=10030676;
RA   White G.R.M., Varley J.M., Heighway J.;
RT   "Isolation and characterisation of a human homologue of the
RT   latrophilin gene from a region of 1p31.1 implicated in breast
RT   cancer.";
RL   Oncogene 17:3513-3519(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20225451; PubMed=10760572;
RA   White G.R.M., Varley J.M., Heighway J.;
RT   "Genomic structure and expression profile of LPHH1, a 7TM gene
RT   variably expressed in breast cancer cell lines.";
RL   Biochim. Biophys. Acta 1491:75-92(2000).
DR   EMBL; AJ244509; CAB60204.1; -.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  1025 MW;  4D59D763344AAB0B CRC64;

  Query Match              33.3%;  Score 17;  DB 4;  Length 8;
  Best Local Similarity   40.0%;  Pred. No. 6.7e+05;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 NAPYL 9
              | |::
Db          4 NKPFI 8


RESULT 3
Q9UM87
ID   Q9UM87      PRELIMINARY;      PRT;     9 AA.
AC   Q9UM87;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   FGFR2 protein (Fragment).
GN   FGFR2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96253074; PubMed=8676562;
RA   Wada C., Ishigaki M., Toyo-oka Y., Yamabe H., Ohnuki Y., Takada F.,
RA   Yamazaki Y., Ohtani H.;
RT   "[Nucleotide sequences at intron 6 and exon 7 junction of fibroblast
RT   growth factor receptor 2 and rapid mutational analysis in Apert
RT   syndrome].";
RL   Rinsho Byori 44:435-438(1996).
DR   EMBL; S82438; AAD14392.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  1103 MW;  9E4D20477401F775 CRC64;

  Query Match              33.3%;  Score 17;  DB 4;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 6.7e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 PNAPYL 9
              |: | |
Db          3 PHRPIL 8
```

```
RESULT 4
Q96T78
ID   Q96T78      PRELIMINARY;      PRT;     9 AA.
AC   Q96T78;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Contactin-associated protein 2 (Fragment).
GN   CNTNAP2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21250995; PubMed=11352571;
RA   Nakabayashi K., Scherer S.W.;
RT   "The human contactin-associated protein 2 (CNTNAP2) spans over 2 Mb of
RT   DNA at chromosome 7q35.";
RL   Genomics 73:108-112(2001).
DR   EMBL; AF318295; AAK49906.1; -.
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1080 MW;  9139A2D5A77B51EA CRC64;

  Query Match              33.3%;  Score 17;  DB 4;  Length 9;
  Best Local Similarity   33.3%;  Pred. No. 6.7e+05;
  Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 FPNAPY 8
              :|:  |
Db          4 YPSVSY 9


RESULT 5
P82429
ID   P82429      PRELIMINARY;      PRT;     9 AA.
AC   P82429;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   44 kDa cell wall protein (Fragment).
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CV. PETIT HAVANA;
RA   Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
RA   Wojtaszek P., Bolwell G.P.;
RT   "Proteomic study of secondary cell wall proteins from transformed
RT   tobacco culture.";
RL   Planta 0:0-0(2000).
CC   -!- SUBCELLULAR LOCATION: CELL WALL.
CC   -!- TISSUE SPECIFICITY: XYLEM.
KW   Cell wall.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  986 MW;  C22CCAADC6C77776 CRC64;

  Query Match              33.3%;  Score 17;  DB 10;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 6.7e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 PNAPYL 9
              | | :|
Db          4 PQADFL 9


RESULT 6
Q9R7T2
ID   Q9R7T2      PRELIMINARY;      PRT;     8 AA.
```

```
AC   Q9R7T2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical 1.0 kDa protein (Fragment).
GN   YQFG.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97061202; PubMed=8905232;
RA   Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA   Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA   Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA   Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA   Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA   Yano M., Horiuchi T.;
RT   "A 718-kb DNA Sequence of Escherichia coli K-12 Genome Corresponding
RT   to the 12.7-28.0 min Region on the Linkage Map.";
RL   DNA Res. 3:137-155(1996).
DR   EMBL; D90705; BAA35310.1; -.
KW   Hypothetical protein.
FT   NON_TER      1      1
SQ   SEQUENCE   8 AA;  964 MW;  DF133B1DD04B476A CRC64;

  Query Match             31.4%;  Score 16;  DB 2;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 6.7e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        7 PYL 9
            ||:
Db        2 PYI 4


RESULT 7
Q9P8E5
ID   Q9P8E5      PRELIMINARY;      PRT;     9 AA.
AC   Q9P8E5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HIS4 protein (Fragment).
GN   HIS4.
OS   Kluyveromyces lactis (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX   NCBI_TaxID=28985;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NRRL-Y1140;
RX   MEDLINE=99448382; PubMed=10518937;
RA   Lamas-Maceiras M., Esperanza Cerdan E., Freire-Picos M.A.;
RT   "Kluyveromyces lacxtis HIS4 transcriptional regulation: similarities
RT   and differences to Saccharomyces cerevisiae HIS4 gene.";
RL   FEBS Lett. 458:72-76(1999).
DR   EMBL; AJ238494; CAB87125.1; -.
FT   NON_TER      9      9
SQ   SEQUENCE   9 AA;  1015 MW;  5770D2D772D2D767 CRC64;

  Query Match             31.4%;  Score 16;  DB 3;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 6.7e+05;
  Matches    3;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        2 MFPNAP 7
            | |  |
Db        1 MLPVVP 6


RESULT 8
Q9TWH6
ID   Q9TWH6      PRELIMINARY;      PRT;     8 AA.
AC   Q9TWH6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   BIOACTIVE peptide P4=PUTATIVE ESOPHAGEAL NEUROREGULATOR.
OS   Perinereis vancaurica.
OC   Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;
OC   Phyllodocida; Nereididae; Perinereis.
OX   NCBI_TaxID=6355;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95323338; PubMed=7599979;
RA   Takahashi T., Furukawa Y., Muneoka Y., Matsushima O., Ikeda T.,
RA   Fujita T., Minakata H., Nomoto K.;
RT   "Isolation and characterization of four novel bioactive peptides from
RT   a polychaete annelid, Perinereis vancaurica.";
RL   Comp. Biochem. Physiol. C,
RL   Pharmacol. Toxicol. Endocrinol. 110:297-304(1995).
SQ   SEQUENCE   8 AA;  989 MW;  954772CAA87B0B59 CRC64;

  Query Match             29.4%;  Score 15;  DB 5;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 6.7e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        5 NAPY 8
            : ||
Db        5 DVPY 8


RESULT 9
Q9H522
ID   Q9H522      PRELIMINARY;      PRT;     9 AA.
AC   Q9H522;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   BA90M5.3 (HCG-1) (Fragment).
GN   BA90M5.3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tromans A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL139188; CAC15103.1; -.
FT   NON_TER      9      9
SQ   SEQUENCE   9 AA;  951 MW;  96A3ADC772C455A5 CRC64;

  Query Match             29.4%;  Score 15;  DB 4;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 6.7e+05;
  Matches    3;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        2 MFPNAP 7
            | |  |
Db        1 MSSNVP 6


RESULT 10
Q9UCS8
ID   Q9UCS8      PRELIMINARY;      PRT;     9 AA.
AC   Q9UCS8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Apolipoprotein A-I (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
```

```
RP   SEQUENCE.
RX   MEDLINE=92075698; PubMed=1742316;
RA   Ehnholm C., Bozas S.E., Tenkanen H., Kirszbaum L., Metso J.,
RA   Murphy B., Walker I.D.;
RT   "The apolipoprotein A-I binding protein of placenta and the SP-40,40
RT   protein of human blood are different proteins which both bind to
RT   apolipoprotein A-I.";
RL   Biochim. Biophys. Acta 1086:255-260(1991).
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  981 MW;  7FE37775A6C7776B CRC64;

  Query Match             29.4%;  Score 15;  DB 4;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 6.7e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        4 PNAP 7
            | :|
Db        4 PQSP 7

RESULT 11
Q9S8J8
ID   Q9S8J8      PRELIMINARY;      PRT;     9 AA.
AC   Q9S8J8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ORYZATENSIN=BIOACTIVE peptide.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95102521; PubMed=7804141;
RA   Takahashi M., Moriguchi S., Yoshikawa M., Sasaki R.;
RT   "Isolation and characterization of oryzatensin: a novel bioactive
RT   peptide with ileum-contracting and immunomodulating activities derived
RT   from rice albumin.";
RL   Biochem. Mol. Biol. Int. 33:1151-1158(1994).
SQ   SEQUENCE   9 AA;  1093 MW;  0E8C67377B56877B CRC64;

  Query Match             29.4%;  Score 15;  DB 10;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 6.7e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        2 MFP 4
            |:|
Db        4 MYP 6

RESULT 12
O98866
ID   O98866      PRELIMINARY;      PRT;     7 AA.
AC   O98866;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   Cytochrome b/f subunit IV (Fragment).
OS   Spinacia oleracea (Spinach).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86120353; PubMed=3003688;
RA   Sijben-Mueller G., Hallick R.B., Alt J., Westhoff P., Herrmann R.G.;
RT   "Spinach plastid genes coding for initiation factor IF-1, ribosomal
RT   protein S11 and RNA polymerase alpha-subunit.";
RL   Nucleic Acids Res. 14:1029-1044(1986).
DR   EMBL; X03496; CAA27215.1; -.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  907 MW;  644729D77409C420 CRC64;

  Query Match             27.5%;  Score 14;  DB 8;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 6.7e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        7 PYL 9
            |:|
Db        4 PFL 6

RESULT 13
Q07624
ID   Q07624      PRELIMINARY;      PRT;     7 AA.
AC   Q07624;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Rous sarcoma virus.
OC   Viruses; Retroid viruses; Retroviridae; Alpharetrovirus.
OX   NCBI_TaxID=11886;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PRAGUE C;
RX   MEDLINE=93010967; PubMed=1327749;
RA   Donze O., Spahr P.F.;
RT   "Role of the open reading frames of Rous sarcoma virus leader RNA in
RT   translation and genome packaging.";
RL   EMBO J. 11:3747-3757(1992).
DR   EMBL; X67587; CAA47862.1; -.
KW   Hypothetical protein.
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  672 MW;  776045A7687DD6F0 CRC64;

  Query Match             27.5%;  Score 14;  DB 15;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 6.7e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        4 PNAP 7
            |: |
Db        4 PSIP 7

RESULT 14
P82858
ID   P82858      PRELIMINARY;      PRT;     8 AA.
AC   P82858;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Uricase (EC 1.7.3.3) (Urate oxidase) (Fragment).
OS   Puccinia recondita f. sp. triseti.
OC   Eukaryota; Fungi; Basidiomycota; Urediniomycetes; Urediniomycetidae;
OC   Uredinales; Pucciniaceae; Puccinia.
OX   NCBI_TaxID=142679;
RN   [1]
RP   SEQUENCE, FUNCTION, AND DEVELOPMENTAL STAGE.
RC   TISSUE=SPORE;
RA   Aguilar M., Montalbini P., Pineda M.;
RL   Submitted (NOV-2000) to the SWISS-PROT data bank.
CC   -!- FUNCTION: INVOLVED IN HOST-PARASITE RELATIONSHIP BETWEEN PLANTS
CC       AND FUNGI.
CC   -!- CATALYTIC ACTIVITY: URATE + O(2) + H(2)O = 5-HYDROXYISOURATE +
CC       H(2)O(2) (5-HYDROXYISOURATE DECOMPOSE TO FORM ALLANTOIN).
CC   -!- SUBCELLULAR LOCATION: PEROXISOMAL (BY SIMILARITY).
CC   -!- DEVELOPMENTAL STAGE: GERMINATION.
CC   -!- SIMILARITY: BELONGS TO THE URICASE FAMILY.
DR   InterPro; IPR002042; Uricase.
```

```
DR   PROSITE; PS00366; URICASE; PARTIAL.
KW   Oxidoreductase; Purine metabolism; Peroxisome.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  777 MW;  98C1ADD735B9D76D CRC64;

  Query Match             27.5%;  Score 14;  DB 3;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 6.7e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        6 APY 8
            ||:
Db        1 APF 3


RESULT 15
Q8WFR5
ID   Q8WFR5      PRELIMINARY;      PRT;     8 AA.
AC   Q8WFR5;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Cytochrome oxidase subunit II (Fragment).
GN   COII.
OS   Diadema paucispinum.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Diadematacea; Diadematoida; Diadematidae;
OC   Diadema.
OX   NCBI_TaxID=145530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HI1;
RX   MEDLINE=21323357; PubMed=11430656;
RA   Lessios H.A., Kessing B.D., Pearse J.S.;
RT   "Population structure and speciation in tropical seas: global
RT   phylogeography of the sea urchin Diadema.";
RL   Evolution 55:955-975(2001).
DR   EMBL; AY012959; AAL33852.1; -.
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   8 AA;  954 MW;  C41B173B46DDC2CE CRC64;

  Query Match             27.5%;  Score 14;  DB 8;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 6.7e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        6 APYL 9
            | ||
Db        3 AQYL 6


Search completed: April 25, 2003, 06:59:12
Job time : 31 secs
```

```
                    GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         April 25, 2003, 06:48:39 ; Search time 34 Seconds
                                           (without alignments)
                                           35.272 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:      130868

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      A_Geneseq_101002:*
                1:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                2:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                3:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                4:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                5:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                6:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                7:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                8:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                9:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                10:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                11:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                12:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                13:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                14:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                15:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                16:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                17:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                18:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                19:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                20:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                21:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                22:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                23:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                 %
Result           Query
  No.   Score    Match Length DB  ID                   Description
---------------------------------------------------------------------------
   1      51    100.0      9  21  AAY94202             Human cytotoxic T
   2      51    100.0      9  21  AAY98670             WT1 derived immuno
   3      51    100.0      9  21  AAY98778             WT1 derived immuno
   4      51    100.0      9  21  AAY80200             Human Wilms' tumou
   5      51    100.0      9  22  AAU68769             Human Wilm's tumou
   6      51    100.0      9  22  AAU68877             Mouse Wilm's tumou
   7      51    100.0      9  22  AAG62002             Human WT1 immunoge
   8      51    100.0      9  22  AAG62110             Mouse WT1 immunoge
   9      51    100.0      9  23  ABG33239             Human WT1 immunoge
  10      51    100.0      9  23  ABG33347             Mouse WT1 immunoge
```

```
  11      51    100.0      9  23  AAE17298             Human leukocyte an
  12      47     92.2      9  21  AAY98523             WT1 derived immuno
  13      47     92.2      9  22  AAU68622             Human Wilm's tumou
  14      47     92.2      9  22  AAG61855             Human WT1 immunoge
  15      47     92.2      9  23  ABG33092             Human WT1 immunoge
  16      46     90.2      9  21  AAY98809             WT1 related peptid
  17      46     90.2      9  22  AAU68908             Wilm's tumour prot
  18      46     90.2      9  22  AAG62141             Human/mouse WT1 im
  19      46     90.2      9  23  ABG33378             Human/mouse WT1 im
  20      41     80.4      9  21  AAY98564             WT1 derived immuno
  21      41     80.4      9  21  AAY98752             WT1 derived immuno
  22      41     80.4      9  22  AAU68663             Human Wilm's tumou
  23      41     80.4      9  22  AAU68851             Mouse Wilm's tumou
  24      41     80.4      9  22  AAG61896             Human WT1 immunoge
  25      41     80.4      9  22  AAG62084             Mouse WT1 immunoge
  26      41     80.4      9  23  ABG33133             Human WT1 immunoge
  27      41     80.4      9  23  ABG33321             Mouse WT1 immunoge
  28      40     78.4      9  21  AAY98808             WT1 related peptid
  29      40     78.4      9  22  AAU68907             Wilm's tumour prot
  30      40     78.4      9  22  AAG62140             Human/mouse WT1 im
  31      35     68.6      9  21  AAY98810             WT1 related peptid
  32      35     68.6      9  22  AAU68909             Wilm's tumour prot
  33      35     68.6      9  22  AAG62142             Human/mouse WT1 im
  34      35     68.6      9  23  ABG33379             Human/mouse WT1 im
  35      33     64.7      9  21  AAY98583             WT1 derived immuno
  36      33     64.7      9  21  AAY98754             WT1 derived immuno
  37      33     64.7      9  22  AAU68682             Human Wilm's tumou
  38      33     64.7      9  22  AAU68853             Mouse Wilm's tumou
  39      33     64.7      9  22  AAG61915             Human WT1 immunoge
  40      33     64.7      9  22  AAG62086             Mouse WT1 immunoge
  41      33     64.7      9  23  ABG33152             Human WT1 immunoge
  42      33     64.7      9  23  ABG33323             Mouse WT1 immunoge
  43      29     56.9      9  21  AAY98697             WT1 derived immuno
  44      29     56.9      9  22  AAU68796             Human Wilm's tumou
  45      29     56.9      9  22  AAE07623             Human PUMP-1 pepti

                                 ALIGNMENTS

RESULT 1
AAY94202
ID   AAY94202 standard; peptide; 9 AA.
XX
AC   AAY94202;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human cytotoxic T lymphocyte-recognised WT1 peptide WT126-34.
XX
KW   WT126-3; peptide; epitope; Wilm's tumour gene; leukaemia;
KW   breast cancer; melanoma; ovarian cancer; immunotherapy.
XX
OS   Homo sapiens.
XX
PN   WO200026249-A1.
XX
PD   11-MAY-2000.
XX
PF   02-NOV-1999;   99WO-GB03572.
XX
PR   02-NOV-1998;   98GB-0023897.
XX
PA   (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI   Stauss HJ,  Gao L;
XX
DR   WPI; 2000-376123/32.
XX
PT   Novel peptides comprising WT-1 and GATA-1 epitopes, their fragments or
PT   variants, useful as vaccines for cancer immunotherapy  -
XX
PS   Claim 1; Page 74; 93pp; English.
```

```
XX
CC   The present sequence is peptide epitope WT126-34, produced by WT1
CC   expressing cells and found at residues 126-134 of the WT1 protein, which
CC   is recognised by cytotoxic T lymphocytes. WT1 is aberrantly expressed in
CC   leukaemias, breast cancer, melanoma and ovarian cancer. The peptide can
CC   be used as a vaccine to stimulate the elimination, by cytotoxic T
CC   lymphocytes, of cancer cells aberrantly expressing WT1. In addition, the
CC   nucleic acid encoding the peptide may also be used in the same manner.
CC   Alternatively, the peptide may be used in vitro to produce activated
CC   cytotoxic T lymphocytes.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db          1 RMFPNAPYL 9


RESULT 2
AAY98670
ID   AAY98670 standard; Peptide; 9 AA.
XX
AC   AAY98670;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:185.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 171; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
```

```
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db          1 RMFPNAPYL 9


RESULT 3
AAY98778
ID   AAY98778 standard; Peptide; 9 AA.
XX
AC   AAY98778;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:293.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 186; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
```

```
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match              100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 RMFPNAPYL 9
               |||||||||
Db         1 RMFPNAPYL 9


RESULT 4
AAY80200
ID   AAY80200 standard; Peptide; 9 AA.
XX
AC   AAY80200;
XX
DT   24-MAY-2000  (first entry)
XX
DE   Human Wilms' tumour suppressor gene WT1 product peptide SEQ ID NO:5.
XX
KW   Wilm's tumour suppressor gene; WT1; cancer; antigen; vaccine; MHC;
KW   major histocompatibility complex; leukaemia; tumour; antitumour.
XX
OS   Homo sapiens.
XX
PN   WO200006602-A1.
XX
PD   10-FEB-2000.
XX
PF   30-JUL-1999;   99WO-JP04130.
XX
PR   31-JUL-1998;   98JP-0218093.
XX
PA   (SUGI/) SUGIYAMA H.
XX
PI   Sugiyama H,  Oka Y;
XX
DR   WPI; 2000-195264/17.
XX
PT   Cancer antigens based on Wilm's tumor suppressor gene WT1 product or
PT   peptide derivatives, for cancer vaccines in treating leukemia and solid
PT   tumors e.g. stomach cancer, skin cancer, lung cancer and breast cancer
XX
PS   Claim 5; Page 18; 48pp; Japanese.
XX
CC   The present invention describes a cancer antigen containing the active
CC   component of Wilms' tumour suppressor gene WT1 product, or partial
CC   peptides, for cancer vaccines in treating leukaemia and solid tumours.
CC   The cancer antigens are useful for cancer vaccines in treating
CC   leukaemia, bone-marrow abnormal formation syndrome, malignant lymphoma,
CC   multiple myeloma, stomach cancer, cancer of the large intestine, lung
CC   cancer, breast cancer, blastoma, liver cancer, skin cancer, bladder
CC   cancer, prostate cancer, uterus cancer, cervical cancer, or ovary
CC   cancer. The present sequence represents a peptide from the human
CC   Wilms' tumour suppressor gene WT1 product.
XX
SQ   Sequence   9 AA;

  Query Match              100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 RMFPNAPYL 9
               |||||||||
Db         1 RMFPNAPYL 9
```

```
RESULT 5
AAU68769
ID   AAU68769 standard; Peptide; 9 AA.
XX
AC   AAU68769;
XX
DT   16-JAN-2002  (first entry)
XX
DE   Human Wilm's tumour protein, WT1, antigenic peptide #164.
XX
KW   Human; Wilms' tumour; WT1; pleural mesothelioma; antigen;
KW   leukaemia; acute myeloid leukaemia; AML; chronic myeloid leukaemia; CML;
KW   acute lymphocytic leukaemia; ALL; myelodysplastic syndromes;
KW   myeloproliferative syndrome; cancer; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200162920-A2.
XX
PD   30-AUG-2001.
XX
PF   22-FEB-2001; 2001WO-US05702.
XX
PR   22-FEB-2000; 2000US-184070P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Cheever MA,  Gaiger A;
XX
DR   WPI; 2001-648218/74.
XX
PT   Composition for the treatment of mesothelioma comprises specific
PT   peptides i.e. Wilms' tumour antigen polypeptide derived antigenic
PT   fragments -
XX
PS   Claim 1; Page 176; 242pp; English.
XX
CC   The invention relates to the use of a composition comprising at least a
CC   first isolated peptide, of between 9 and 40 amino acids or a first
CC   nucleic acid, encoding the peptide, in the manufacture of a medicament
CC   for treating or preventing mesothelioma. The peptides are antigenic
CC   peptides derived from the Wilms' tumour protein WT1. The composition is
CC   useful for the treatment of mesothelioma, Wilms' tumour, preferably
CC   pleural mesothelioma and other WT1 associated malignancies e.g.
CC   leukaemia (including acute myeloid leukaemia, AML, chronic myeloid
CC   leukaemia, CML, acute lymphocytic leukaemia, ALL, and childhood ALL),
CC   myelodysplastic syndromes, myeloproliferative syndromes and cancers
CC   (e.g. breast, testicular, prostate, lung and ovarian) in mammals,
CC   preferably humans. The present sequence is an antigenic peptide of
CC   the invention derived from human WT1.
XX
SQ   Sequence   9 AA;

  Query Match              100.0%;  Score 51;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 RMFPNAPYL 9
               |||||||||
Db         1 RMFPNAPYL 9


RESULT 6
AAU68877
ID   AAU68877 standard; Peptide; 9 AA.
XX
AC   AAU68877;
XX
DT   16-JAN-2002  (first entry)
XX
DE   Mouse Wilm's tumour protein, WT1, antigenic peptide #46.
XX
KW   Mouse; Wilms' tumour; WT1; pleural mesothelioma; antigen;
```

```
KW   leukaemia; acute myeloid leukaemia; AML; chronic myeloid leukaemia; CML;
KW   acute lymphocytic leukaemia; ALL; myelodysplastic syndromes;
KW   myeloproliferative syndrome; cancer; cytostatic.
XX
OS   Mus musculus.
XX
PN   WO200162920-A2.
XX
PD   30-AUG-2001.
XX
PF   22-FEB-2001; 2001WO-US05702.
XX
PR   22-FEB-2000; 2000US-184070P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Cheever MA,  Gaiger A;
XX
DR   WPI; 2001-648218/74.
XX
PT   Composition for the treatment of mesothelioma comprises specific
PT   peptides i.e. Wilms' tumour antigen polypeptide derived antigenic
PT   fragments -
XX
PS   Claim 1; Page 24; 242pp; English.
XX
CC   The invention relates to the use of a composition comprising at least a
CC   first isolated peptide, of between 9 and 40 amino acids or a first
CC   nucleic acid, encoding the peptide, in the manufacture of a medicament
CC   for treating or preventing mesothelioma. The peptides are antigenic
CC   peptides derived from the Wilms' tumour protein WT1. The composition is
CC   useful for the treatment of mesothelioma, Wilms' tumour, preferably
CC   pleural mesothelioma and other WT1 associated malignancies e.g.
CC   leukaemia (including acute myeloid leukaemia, AML, chronic myeloid
CC   leukaemia, CML, acute lymphocytic leukaemia, ALL, and childhood ALL),
CC   myelodysplastic syndromes, myeloproliferative syndromes and cancers
CC   (e.g. breast, testicular, prostate, lung and ovarian) in mammals,
CC   preferably humans. The present sequence is an antigenic peptide of
CC   the invention derived from mouse WT1.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db          1 RMFPNAPYL 9


RESULT 7
AAG62002
ID   AAG62002 standard; Peptide; 9 AA.
XX
AC   AAG62002;
XX
DT   06-JUL-2001  (first entry)
XX
DE   Human WT1 immunogenic peptide SEQ ID NO: 185.
XX
KW   Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;
KW   chromosome 11p13; zinc finger transcription factor.
XX
OS   Homo sapiens.
XX
PN   WO200125273-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27465.
XX
PR   04-OCT-1999;   99US-0157459.
```

```
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Xu J,  Cheever MA,  Reed SG;
XX
DR   WPI; 2001-328324/34.
XX
PT   Polypeptide comprising part of the Wilm's Tumour gene product sequence is
PT   used in the diagnosis and treatment of malignant diseases e.g. leukemia
PT   and cancer associated with WTI  -
XX
PS   Claim 4; Page 183; 228pp; English.
XX
CC   The present invention describes compositions comprising peptides derived
CC   from the Wilm's tumour protein WT1 and methods for their use in treating
CC   malignant diseases. Peptides derived from both the murine and human WT1
CC   proteins are provided. The human WT1 gene is found on chromosome 11p13,
CC   and the protein was shown to be a zinc finger transcription factor. The
CC   immunogenic peptides of the invention are particularly useful in the
CC   diagnosis and treatment of cancer and leukaemia. The present sequence is
CC   a polypeptide described in the exemplification of the invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db          1 RMFPNAPYL 9


RESULT 8
AAG62110
ID   AAG62110 standard; Peptide; 9 AA.
XX
AC   AAG62110;
XX
DT   06-JUL-2001  (first entry)
XX
DE   Mouse WT1 immunogenic peptide SEQ ID NO: 293.
XX
KW   Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;
KW   chromosome 11p13; zinc finger transcription factor.
XX
OS   Mus musculus.
XX
PN   WO200125273-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27465.
XX
PR   04-OCT-1999;   99US-0157459.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Xu J,  Cheever MA,  Reed SG;
XX
DR   WPI; 2001-328324/34.
XX
PT   Polypeptide comprising part of the Wilm's Tumour gene product sequence is
PT   used in the diagnosis and treatment of malignant diseases e.g. leukemia
PT   and cancer associated with WTI  -
XX
PS   Claim 4; Page 200; 228pp; English.
XX
CC   The present invention describes compositions comprising peptides derived
CC   from the Wilm's tumour protein WT1 and methods for their use in treating
CC   malignant diseases. Peptides derived from both the murine and human WT1
CC   proteins are provided. The human WT1 gene is found on chromosome 11p13,
CC   and the protein was shown to be a zinc finger transcription factor. The
```

```
CC   immunogenic peptides of the invention are particularly useful in the
CC   diagnosis and treatment of cancer and leukaemia. The present sequence is
CC   a polypeptide described in the exemplification of the invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db          1 RMFPNAPYL 9


RESULT 9
ABG33239
ID   ABG33239 standard; Peptide; 9 AA.
XX
AC   ABG33239;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human WT1 immunogenic peptide #162.
XX
KW   Human; mouse; cytostatic; immunostimulant;  WT1; cancer;
KW   immune response.
XX
OS   Homo sapiens.
XX
PN   WO200228414-A1.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US31139.
XX
PR   06-OCT-2000; 2000US-0684361.
PR   09-OCT-2000; 2000US-0685830.
PR   15-FEB-2001; 2001US-0785019.
PR   24-AUG-2001; 2001US-0938864.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  McNeill PD,  Smithgall M,  Moulton G,  Vedvick TS;
PI   Sleath PR,  Mossman S,  Evans L,  Spies AG,  Boydston J;
XX
DR   WPI; 2002-352217/38.
XX
PT   Novel isolated WT1 polynucleotide, and encoded polypeptide, useful for
PT   treating and diagnosing cancer in a patient -
XX
PS   Example 4; Page 194; 260pp; English.
XX
CC   The invention relates to an isolated WT1 polynucleotide (I) and
CC   polypeptide encoded by (I). The WT1 polynucleotides and polypeptides
CC   are used for treating and detecting cancer in a patient, and for
CC   stimulating an immune response in patient. ABG33070-ABG33405
CC   represent WT1 amino acid sequences of the invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 23;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db          1 RMFPNAPYL 9


RESULT 10
ABG33347
ID   ABG33347 standard; Peptide; 9 AA.
XX
AC   ABG33347;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Mouse WT1 immunogenic peptide #44.
XX
KW   Human; mouse; cytostatic; immunostimulant;  WT1; cancer;
KW   immune response.
XX
OS   Mus musculus.
XX
PN   WO200228414-A1.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US31139.
XX
PR   06-OCT-2000; 2000US-0684361.
PR   09-OCT-2000; 2000US-0685830.
PR   15-FEB-2001; 2001US-0785019.
PR   24-AUG-2001; 2001US-0938864.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  McNeill PD,  Smithgall M,  Moulton G,  Vedvick TS;
PI   Sleath PR,  Mossman S,  Evans L,  Spies AG,  Boydston J;
XX
DR   WPI; 2002-352217/38.
XX
PT   Novel isolated WT1 polynucleotide, and encoded polypeptide, useful for
PT   treating and diagnosing cancer in a patient -
XX
PS   Example 4; Page 210; 260pp; English.
XX
CC   The invention relates to an isolated WT1 polynucleotide (I) and
CC   polypeptide encoded by (I). The WT1 polynucleotides and polypeptides
CC   are used for treating and detecting cancer in a patient, and for
CC   stimulating an immune response in patient. ABG33070-ABG33405
CC   represent WT1 amino acid sequences of the invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 23;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db          1 RMFPNAPYL 9


RESULT 11
AAE17298
ID   AAE17298 standard; peptide; 9 AA.
XX
AC   AAE17298;
XX
DT   18-APR-2002  (first entry)
XX
DE   Human leukocyte antigen (HLA-A2.1) restricted peptide, Db126.
XX
KW   Human; artificial antigen presenting cell; AAPC; beta2-microglobulin;
KW   human leukocyte antigen; HLA; major histocompatibility complex; MHC;
KW   cytotoxic T lymphocyte; CTL; T cell-specific antigen; TCA; antitumour;
KW   immune response; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200194944-A2.
XX
```

```
PD   13-DEC-2001.
XX
PF   01-JUN-2001; 2001WO-US17981.
XX
PR   02-JUN-2000; 2000US-209157P.
XX
PA   (SLOK ) SLOAN KETTERING INST CANCER RES.
XX
PI   Sadelain M,  Latouche J;
XX
DR   WPI; 2002-139667/18.
XX
PT   Artificial antigen presenting cells for activating T lymphocytes,
PT   comprises eukaryotic cell expressing antigen presenting complex having
PT   beta2-microglobulin, exogenous accessory molecule, human leukocyte
PT   antigen molecule and protein -
XX
PS   Example 17; Page 40; 75pp; English.
XX
CC   The present invention relates to an artificial antigen presenting cell
CC   (AAPC) comprising a eukaryotic cell expressing an antigen presenting
CC   complex comprising beta2-microglobulin, an exogenous accessory molecule,
CC   a human leukocyte antigen, HLA (major histocompatibility complex, MHC)
CC   molecule of a single type and a protein that is processed intracellularly
CC   to produce an exogenous T cell-specific epitope. The invention also
CC   relates to methods for activation of T lymphocytes. The method is also
CC   useful for identifying within a test population of cytotoxic T
CC   lymphocytes (CTLs), CTLs specifically activated against a known T-cell
CC   specific antigen (TCA), which is useful for diagnostic purposes. AAPC is
CC   also useful for activating CTLs, by contacting AAPC with a suitable
CC   population of T lymphocytes under conditions suitable for the activation
CC   and isolating the activated CTLs. AAPC is further useful for the
CC   investigation of primary T cell activation and diagnostic applications
CC   here primary T cell activation allow discovery of antigens and accessory
CC   molecules, and diagnostic applications include cell-based assays for
CC   quantifying immune response in normal, infected or treated (vaccinated)
CC   patients. Composition comprising AAPC or activated T cells produced by
CC   utilising AAPC is useful for eliciting an antitumour response. The
CC   invention is used for the treatment of cancer. The present sequence is
CC   human HLA-A2.1 restricted peptide used in additional AAPC-induced CTL-
CC   activation.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 23;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
                |||||||||
Db          1 RMFPNAPYL 9


RESULT 12
AAY98523
ID   AAY98523 standard; Peptide; 9 AA.
XX
AC   AAY98523;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:38.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
```

```
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 149; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match              92.2%;  Score 47;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPY 8
                ||||||||
Db          2 RMFPNAPY 9


RESULT 13
AAU68622
ID   AAU68622 standard; Peptide; 9 AA.
XX
AC   AAU68622;
XX
DT   16-JAN-2002  (first entry)
XX
DE   Human Wilm's tumour protein, WT1, antigenic peptide #17.
XX
KW   Human; Wilms' tumour; WT1; pleural mesothelioma; antigen;
KW   leukaemia; acute myeloid leukaemia; AML; chronic myeloid leukaemia; CML;
KW   acute lymphocytic leukaemia; ALL; myelodysplastic syndromes;
KW   myeloproliferative syndrome; cancer; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200162920-A2.
XX
PD   30-AUG-2001.
XX
PF   22-FEB-2001; 2001WO-US05702.
XX
```

```
PR   22-FEB-2000; 2000US-184070P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Cheever MA,  Gaiger A;
XX
DR   WPI; 2001-648218/74.
XX
PT   Composition for the treatment of mesothelioma comprises specific
PT   peptides i.e. Wilms' tumour antigen polypeptide derived antigenic
PT   fragments -
XX
PS   Claim 1; Page 209; 242pp; English.
XX
CC   The invention relates to the use of a composition comprising at least a
CC   first isolated peptide, of between 9 and 40 amino acids or a first
CC   nucleic acid, encoding the peptide, in the manufacture of a medicament
CC   for treating or preventing mesothelioma. The peptides are antigenic
CC   peptides derived from the Wilms' tumour protein WT1. The composition is
CC   useful for the treatment of mesothelioma, Wilms' tumour, preferably
CC   pleural mesothelioma and other WT1 associated malignancies e.g.
CC   leukaemia (including acute myeloid leukaemia, AML, chronic myeloid
CC   leukaemia, CML, acute lymphocytic leukaemia, ALL, and childhood ALL),
CC   myelodysplastic syndromes, myeloproliferative syndromes and cancers
CC   (e.g. breast, testicular, prostate, lung and ovarian) in mammals,
CC   preferably humans. The present sequence is an antigenic peptide of
CC   the invention derived from human WT1.
XX
SQ   Sequence   9 AA;

  Query Match             92.2%;  Score 47;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPY 8
          ||||||||
Db      2 RMFPNAPY 9


RESULT 14
AAG61855
ID   AAG61855 standard; Peptide; 9 AA.
XX
AC   AAG61855;
XX
DT   06-JUL-2001  (first entry)
XX
DE   Human WT1 immunogenic peptide SEQ ID NO: 38.
XX
KW   Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;
KW   chromosome 11p13; zinc finger transcription factor.
XX
OS   Homo sapiens.
XX
PN   WO200125273-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27465.
XX
PR   04-OCT-1999;   99US-0157459.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Xu J,  Cheever MA,  Reed SG;
XX
DR   WPI; 2001-328324/34.
XX
PT   Polypeptide comprising part of the Wilm's Tumour gene product sequence is
PT   used in the diagnosis and treatment of malignant diseases e.g. leukemia
PT   and cancer associated with WTI  -
XX
PS   Claim 4; Page 161; 228pp; English.
XX
CC   The present invention describes compositions comprising peptides derived
CC   from the Wilm's tumour protein WT1 and methods for their use in treating
CC   malignant diseases. Peptides derived from both the murine and human WT1
CC   proteins are provided. The human WT1 gene is found on chromosome 11p13,
CC   and the protein was shown to be a zinc finger transcription factor. The
CC   immunogenic peptides of the invention are particularly useful in the
CC   diagnosis and treatment of cancer and leukaemia. The present sequence is
CC   a polypeptide described in the exemplification of the invention.
XX
SQ   Sequence   9 AA;

  Query Match             92.2%;  Score 47;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPY 8
          ||||||||
Db      2 RMFPNAPY 9


RESULT 15
ABG33092
ID   ABG33092 standard; Peptide; 9 AA.
XX
AC   ABG33092;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human WT1 immunogenic peptide #15.
XX
KW   Human; mouse; cytostatic; immunostimulant;  WT1; cancer;
KW   immune response.
XX
OS   Homo sapiens.
XX
PN   WO200228414-A1.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US31139.
XX
PR   06-OCT-2000; 2000US-0684361.
PR   09-OCT-2000; 2000US-0685830.
PR   15-FEB-2001; 2001US-0785019.
PR   24-AUG-2001; 2001US-0938864.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  McNeill PD,  Smithgall M,  Moulton G,  Vedvick TS;
PI   Sleath PR,  Mossman S,  Evans L,  Spies AG,  Boydston J;
XX
DR   WPI; 2002-352217/38.
XX
PT   Novel isolated WT1 polynucleotide, and encoded polypeptide, useful for
PT   treating and diagnosing cancer in a patient -
XX
PS   Example 4; Page 171; 260pp; English.
XX
CC   The invention relates to an isolated WT1 polynucleotide (I) and
CC   polypeptide encoded by (I). The WT1 polynucleotides and polypeptides
CC   are used for treating and detecting cancer in a patient, and for
CC   stimulating an immune response in patient. ABG33070-ABG33405
CC   represent WT1 amino acid sequences of the invention.
XX
SQ   Sequence   9 AA;

  Query Match             92.2%;  Score 47;  DB 23;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPY 8
```

```
              ||||||||
Db          2 RMFPNAPY 9


Search completed: April 25, 2003, 06:57:43
Job time : 35 secs
```

```
                        GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         April 25, 2003, 06:58:40 ; Search time 15 Seconds
                                           (without alignments)
                                           48.078 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters:      29135

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
                1:  /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                2:  /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                3:  /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                4:  /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                5:  /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                6:  /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                7:  /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                8:  /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                9:  /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                10:  /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
                11:  /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                12:  /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
                13:  /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                14:  /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                    Description
-----------------------------------------------------------------------------
    1      51   100.0      9  9  US-10-125-635A-185     Sequence 185, App
    2      51   100.0      9  9  US-10-125-635A-293     Sequence 293, App
    3      51   100.0      9  9  US-09-938-864-185      Sequence 185, App
    4      51   100.0      9  9  US-09-938-864-293      Sequence 293, App
    5      51   100.0      9 10   US-09-872-832-46       Sequence 46, Appl
    6      47    92.2      9  9  US-10-125-635A-38      Sequence 38, Appl
    7      47    92.2      9  9  US-09-938-864-38       Sequence 38, Appl
    8      46    90.2      9  9  US-10-125-635A-324     Sequence 324, App
    9      46    90.2      9  9  US-09-938-864-324      Sequence 324, App
   10      41    80.4      9  9  US-10-125-635A-79      Sequence 79, Appl
   11      41    80.4      9  9  US-10-125-635A-267     Sequence 267, App
   12      41    80.4      9  9  US-09-938-864-79       Sequence 79, Appl
   13      41    80.4      9  9  US-09-938-864-267      Sequence 267, App
   14      40    78.4      9  9  US-10-125-635A-323     Sequence 323, App
   15      40    78.4      9  9  US-09-938-864-323      Sequence 323, App
   16      35    68.6      9  9  US-10-125-635A-325     Sequence 325, App
   17      35    68.6      9  9  US-09-938-864-325      Sequence 325, App
   18      33    64.7      9  9  US-10-125-635A-98      Sequence 98, Appl
   19      33    64.7      9  9  US-10-125-635A-269     Sequence 269, App
   20      33    64.7      9  9  US-09-938-864-98       Sequence 98, Appl
   21      33    64.7      9  9  US-09-938-864-269      Sequence 269, App
   22      29    56.9      9  9  US-09-835-948-158      Sequence 158, App
   23      29    56.9      9  9  US-10-125-635A-212     Sequence 212, App
   24      29    56.9      9  9  US-09-938-864-212      Sequence 212, App
   25      28    54.9      9  9  US-10-125-635A-144     Sequence 144, App
   26      28    54.9      9  9  US-10-125-635A-282     Sequence 282, App
   27      28    54.9      9  9  US-09-938-864-144      Sequence 144, App
   28      28    54.9      9  9  US-09-938-864-282      Sequence 282, App
   29      27    52.9      8 10   US-09-839-666-6        Sequence 6, Appli
   30      26    51.0      8 12   US-10-011-436-10       Sequence 10, Appl
   31      25    49.0      9 10   US-09-884-441-436      Sequence 436, App
   32      25    49.0      9 10   US-09-884-441-447      Sequence 447, App
   33      23    45.1      9  9  US-10-125-635A-322     Sequence 322, App
   34      23    45.1      9  9  US-09-938-864-322      Sequence 322, App
   35      22    43.1      7  9  US-10-012-542-321      Sequence 321, App
   36      22    43.1      7 10   US-09-731-242A-28      Sequence 28, Appl
   37      22    43.1      8  9  US-09-864-011A-8       Sequence 8, Appli
   38      22    43.1      8 10   US-09-863-971A-8       Sequence 8, Appli
   39      22    43.1      9  9  US-09-835-948-180      Sequence 180, App
   40      22    43.1      9  9  US-10-125-635A-326     Sequence 326, App
   41      22    43.1      9  9  US-10-012-542-386      Sequence 386, App
   42      22    43.1      9  9  US-09-938-864-326      Sequence 326, App
   43    21.5    42.2      8  9  US-09-990-574-3        Sequence 3, Appli
   44      21    41.2      7 10   US-09-822-270-9        Sequence 9, Appli
   45      21    41.2      9  9  US-09-835-948-74       Sequence 74, Appl

                              ALIGNMENTS


RESULT 1
US-10-125-635A-185
; Sequence 185, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-125-635A-185

  Query Match             100.0%;  Score 51;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9


RESULT 2
US-10-125-635A-293
; Sequence 293, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
```

```
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-125-635A-293

  Query Match             100.0%;  Score 51;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9


RESULT 3
US-09-938-864-185
; Sequence 185, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-185

  Query Match             100.0%;  Score 51;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9


RESULT 4
US-09-938-864-293
; Sequence 293, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
```

```
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-293

  Query Match             100.0%;  Score 51;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9


RESULT 5
US-09-872-832-46
; Sequence 46, Application US/09872832
; Patent No. US20020131960A1
; GENERAL INFORMATION:
;  APPLICANT: Memorial Sloan-Kettering Cancer Center
;  TITLE OF INVENTION: ARTIFICIAL ANTIGEN PRESENTING CELLS AND METHODS OF USE THEREOF
;  FILE REFERENCE: 830002-2003.1
;  CURRENT APPLICATION NUMBER: US/09/872,832
;  CURRENT FILING DATE:  2001-06-01
;  PRIOR APPLICATION NUMBER: 60/209,157
;  PRIOR FILING DATE: 2000-02-06
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 46
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-872-832-46

  Query Match             100.0%;  Score 51;  DB 10;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db       1 RMFPNAPYL 9


RESULT 6
US-10-125-635A-38
; Sequence 38, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 38
```

```
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-125-635A-38

  Query Match             92.2%;  Score 47;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPY 8
           ||||||||
Db       2 RMFPNAPY 9


RESULT 7
US-09-938-864-38
; Sequence 38, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 38
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-38

  Query Match             92.2%;  Score 47;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPY 8
           ||||||||
Db       2 RMFPNAPY 9


RESULT 8
US-10-125-635A-324
; Sequence 324, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 324
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien and Mus musculus
US-10-125-635A-324

  Query Match             90.2%;  Score 46;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 MFPNAPYL 9
           ||||||||
Db       1 MFPNAPYL 8


RESULT 9
US-09-938-864-324
; Sequence 324, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 324
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien and Mus musculus
US-09-938-864-324

  Query Match             90.2%;  Score 46;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 MFPNAPYL 9
           ||||||||
Db       1 MFPNAPYL 8


RESULT 10
US-10-125-635A-79
; Sequence 79, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 79
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-125-635A-79

  Query Match             80.4%;  Score 41;  DB 9;  Length 9;
```

```
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 FPNAPYL 9
          |||||||
Db      1 FPNAPYL 7

RESULT 11
US-10-125-635A-267
; Sequence 267, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 267
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-125-635A-267

  Query Match            80.4%;  Score 41;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 FPNAPYL 9
          |||||||
Db      1 FPNAPYL 7

RESULT 12
US-09-938-864-79
; Sequence 79, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 79
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-79

  Query Match            80.4%;  Score 41;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      3 FPNAPYL 9
          |||||||
Db      1 FPNAPYL 7

RESULT 13
US-09-938-864-267
; Sequence 267, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 267
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-267

  Query Match            80.4%;  Score 41;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 FPNAPYL 9
          |||||||
Db      1 FPNAPYL 7

RESULT 14
US-10-125-635A-323
; Sequence 323, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 323
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien and Mus musculus
US-10-125-635A-323

  Query Match            78.4%;  Score 40;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAP 7
          |||||||
Db      3 RMFPNAP 9
```

```
RESULT 15
US-09-938-864-323
; Sequence 323, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 323
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien and Mus musculus
US-09-938-864-323

  Query Match             78.4%;  Score 40;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAP 7
            |||||||
Db        3 RMFPNAP 9

Search completed: April 25, 2003, 07:05:57
Job time : 15 secs
```

```
                         GenCore version 5.1.4_p5_4578
                  Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         April 25, 2003, 06:46:25 ; Search time 29 Seconds
                                           (without alignments)
                                           9.131 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters:      262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                 %
Result          Query
  No.   Score   Match Length DB  ID                           Description
 -----------------------------------------------------------------------------
     1     51   100.0    154  1  US-08-102-942A-5              Sequence 5, Appli
     2     51   100.0    154  4  US-09-037-179B-5              Sequence 5, Appli
     3     51   100.0    210  1  US-08-234-783-2               Sequence 2, Appli
     4     51   100.0    210  1  US-08-456-907-2               Sequence 2, Appli
     5     51   100.0    210  5  PCT-US95-05523-2              Sequence 2, Appli
     6     51   100.0    345  1  US-08-102-942A-2              Sequence 2, Appli
     7     51   100.0    345  4  US-09-037-179B-2              Sequence 2, Appli
     8     51   100.0    429  1  US-08-234-783-4               Sequence 4, Appli
     9     51   100.0    429  1  US-08-456-907-4               Sequence 4, Appli
    10     51   100.0    429  5  PCT-US95-05523-4              Sequence 4, Appli
    11     51   100.0    449  1  US-08-102-942A-4              Sequence 4, Appli
    12     51   100.0    449  1  US-08-102-942A-6              Sequence 6, Appli
    13     51   100.0    449  4  US-09-037-179B-4              Sequence 4, Appli
    14     51   100.0    449  4  US-09-037-179B-6              Sequence 6, Appli
    15     37    72.5     50  4  US-08-975-080-26              Sequence 26, Appl
    16     37    72.5    204  4  US-09-134-001C-2975           Sequence 2975, Ap
    17     37    72.5   1151  3  US-08-836-134-23              Sequence 23, Appl
    18     37    72.5   1151  4  US-09-493-784-23              Sequence 23, Appl
    19     37    72.5   1232  3  US-08-836-134-2               Sequence 2, Appli
    20     37    72.5   1232  4  US-09-493-784-2               Sequence 2, Appli
    21     36    70.6     67  2  US-08-511-485-23              Sequence 23, Appl
    22     36    70.6    217  4  US-08-297-431B-33             Sequence 33, Appl
    23     36    70.6    370  3  US-09-150-133-1               Sequence 1, Appli
    24     36    70.6    370  3  US-09-150-133-3               Sequence 3, Appli
    25     36    70.6    370  3  US-09-150-141-1               Sequence 1, Appli
    26     36    70.6    370  3  US-09-150-141-3               Sequence 3, Appli
    27     36    70.6    370  4  US-09-374-493-1               Sequence 1, Appli
    28     36    70.6    370  4  US-09-374-493-3               Sequence 3, Appli
    29     36    70.6    370  4  US-09-374-824-1               Sequence 1, Appli
    30     36    70.6    370  4  US-09-374-824-3               Sequence 3, Appli
    31     36    70.6    370  4  US-09-374-492-1               Sequence 1, Appli
    32     36    70.6    370  4  US-09-374-492-3               Sequence 3, Appli
    33     36    70.6    389  4  US-08-430-286A-7              Sequence 7, Appli
    34     36    70.6    438  5  PCT-US95-05922A-2             Sequence 2, Appli
    35     36    70.6    618  2  US-08-511-485-8               Sequence 8, Appli
    36     36    70.6    618  3  US-09-212-971-8               Sequence 8, Appli
    37     36    70.6    618  4  US-08-800-929A-8              Sequence 8, Appli
    38     36    70.6    618  4  US-08-569-749-2               Sequence 2, Appli
    39     36    70.6    618  4  US-09-617-053A-8              Sequence 8, Appli
    40     36    70.6    618  4  US-09-069-023-29              Sequence 29, Appl
    41     36    70.6    618  5  PCT-US96-12860-2              Sequence 2, Appli
    42     35    68.6    380  3  US-09-150-133-9               Sequence 9, Appli
    43     35    68.6    380  3  US-09-150-141-9               Sequence 9, Appli
    44     35    68.6    380  4  US-09-374-493-9               Sequence 9, Appli
    45     35    68.6    380  4  US-09-374-824-9               Sequence 9, Appli

                                  ALIGNMENTS

RESULT 1
US-08-102-942A-5
; Sequence 5, Application US/08102942A
; Patent No. 5726288
;  GENERAL INFORMATION:
;    APPLICANT:  Call, Katherine M.
;    APPLICANT:  Glaser, Thomas M.
;    APPLICANT:  Ito, Caryn Y.
;    APPLICANT:  Buckler, Alan J.
;    APPLICANT:  Pelletier, Jerry
;    APPLICANT:  Haber, Daniel A.
;    APPLICANT:  Rose, Elise A.
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Bruening, Wendy
;    APPLICANT:  Darveau, Andre
;    TITLE OF INVENTION:  Localization and Characterization of the
;    TITLE OF INVENTION:  Wilms' Tumor Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,942A
;      FILING DATE:  02-AUG-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5194A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  154 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-102-942A-5
```

```
  Query Match             100.0%;  Score 51;  DB 1;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 0.04;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db     40 RMFPNAPYL 48


RESULT 2
US-09-037-179B-5
; Sequence 5, Application US/09037179B
; Patent No. 6316599
; GENERAL INFORMATION:
;  APPLICANT: Call, Katherine M.
;  APPLICANT:  Glaser, Thomas M.
;  APPLICANT:  Ito, Caryn Y.
;  APPLICANT:  Buckler, Alan J.
;  APPLICANT:  Pelletier, Jerry
;  APPLICANT:  Haber, Daniel A.
;  APPLICANT:  Rose, Elise A.
;  APPLICANT:  Housman, David E.
;  APPLICANT:  Bruening, Wendy
;  APPLICANT:  Darveau, Andre
;  TITLE OF INVENTION: Localization and Characterization of the
;  TITLE OF INVENTION:  Wilms' Tumor Gene
;  FILE REFERENCE: 0050.1312-011
;  CURRENT APPLICATION NUMBER: US/09/037,179B
;  CURRENT FILING DATE:  1998-03-09
;  PRIOR APPLICATION NUMBER: US 08/102,942
;  PRIOR FILING DATE: 1993-08-02
;  PRIOR APPLICATION NUMBER: US 07/614,161
;  PRIOR FILING DATE: 1990-11-13
;  PRIOR APPLICATION NUMBER: US 07/435,780
;  PRIOR FILING DATE: 1989-11-13
;  PRIOR APPLICATION NUMBER: US 07/795,323
;  PRIOR FILING DATE: 1994-09-27
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 154
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-037-179B-5

  Query Match             100.0%;  Score 51;  DB 4;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 0.04;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db     40 RMFPNAPYL 48


RESULT 3
US-08-234-783-2
; Sequence 2, Application US/08234783
; Patent No. 5622835
;  GENERAL INFORMATION:
;    APPLICANT:  Herlyn, Meenhard
;    APPLICANT:  Morris, Jennifer
;    APPLICANT:  Rauscher III, Frank J.
;    APPLICANT:  Rodeck, Ulrich
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and Methods of
;    TITLE OF INVENTION:  Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/234,783
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48USA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  210 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-234-783-2

  Query Match             100.0%;  Score 51;  DB 1;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.055;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db    137 RMFPNAPYL 145


RESULT 4
US-08-456-907-2
; Sequence 2, Application US/08456907
; Patent No. 5633142
;  GENERAL INFORMATION:
;    APPLICANT:  Herlyn, Meenhard
;    APPLICANT:  Morris, Jennifer
;    APPLICANT:  Rauscher III, Frank J.
;    APPLICANT:  Rodeck, Ulrich
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and Methods of
;    TITLE OF INVENTION:  Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/456,907
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/234,783
;      FILING DATE:  28-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48AUSA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  210 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-456-907-2

  Query Match             100.0%;  Score 51;  DB 1;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.055;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db      137 RMFPNAPYL 145


RESULT 5
PCT-US95-05523-2
; Sequence 2, Application PC/TUS9505523
;  GENERAL INFORMATION:
;    APPLICANT:  The Wistar Institute of Anatomy and Biology
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and
;    TITLE OF INVENTION:  Methods of Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05523
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/234,783
;      FILING DATE:  28-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  210 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-05523-2

  Query Match             100.0%;  Score 51;  DB 5;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 0.055;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db      137 RMFPNAPYL 145


RESULT 6
US-08-102-942A-2
; Sequence 2, Application US/08102942A
; Patent No. 5726288
;  GENERAL INFORMATION:
;    APPLICANT:  Call, Katherine M.
;    APPLICANT:  Glaser, Thomas M.
;    APPLICANT:  Ito, Caryn Y.
;    APPLICANT:  Buckler, Alan J.
;    APPLICANT:  Pelletier, Jerry
;    APPLICANT:  Haber, Daniel A.
;    APPLICANT:  Rose, Elise A.
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Bruening, Wendy
;    APPLICANT:  Darveau, Andre
;    TITLE OF INVENTION:  Localization and Characterization of the
;    TITLE OF INVENTION:  Wilms' Tumor Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,942A
;      FILING DATE:  02-AUG-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5194A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  345 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-102-942A-2

  Query Match             100.0%;  Score 51;  DB 1;  Length 345;
  Best Local Similarity   100.0%;  Pred. No. 0.093;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db       42 RMFPNAPYL 50


RESULT 7
US-09-037-179B-2
; Sequence 2, Application US/09037179B
; Patent No. 6316599
; GENERAL INFORMATION:
;  APPLICANT: Call, Katherine M.
;  APPLICANT:  Glaser, Thomas M.
;  APPLICANT:  Ito, Caryn Y.
;  APPLICANT:  Buckler, Alan J.
;  APPLICANT:  Pelletier, Jerry
;  APPLICANT:  Haber, Daniel A.
;  APPLICANT:  Rose, Elise A.
;  APPLICANT:  Housman, David E.
;  APPLICANT:  Bruening, Wendy
;  APPLICANT:  Darveau, Andre
;  TITLE OF INVENTION: Localization and Characterization of the
;  TITLE OF INVENTION:  Wilms' Tumor Gene
;  FILE REFERENCE: 0050.1312-011
;  CURRENT APPLICATION NUMBER: US/09/037,179B
;  CURRENT FILING DATE:  1998-03-09
;  PRIOR APPLICATION NUMBER: US 08/102,942
;  PRIOR FILING DATE: 1993-08-02
```

```
;  PRIOR APPLICATION NUMBER: US 07/614,161
;  PRIOR FILING DATE: 1990-11-13
;  PRIOR APPLICATION NUMBER: US 07/435,780
;  PRIOR FILING DATE: 1989-11-13
;  PRIOR APPLICATION NUMBER: US 07/795,323
;  PRIOR FILING DATE: 1994-09-27
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 345
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-037-179B-2

  Query Match             100.0%;  Score 51;  DB 4;  Length 345;
  Best Local Similarity   100.0%;  Pred. No. 0.093;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db       42 RMFPNAPYL 50


RESULT 8
US-08-234-783-4
; Sequence 4, Application US/08234783
; Patent No. 5622835
;  GENERAL INFORMATION:
;    APPLICANT:  Herlyn, Meenhard
;    APPLICANT:  Morris, Jennifer
;    APPLICANT:  Rauscher III, Frank J.
;    APPLICANT:  Rodeck, Ulrich
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and Methods of
;    TITLE OF INVENTION:  Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/234,783
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48USA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-234-783-4

  Query Match             100.0%;  Score 51;  DB 1;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db      126 RMFPNAPYL 134
```

```
RESULT 9
US-08-456-907-4
; Sequence 4, Application US/08456907
; Patent No. 5633142
;  GENERAL INFORMATION:
;    APPLICANT:  Herlyn, Meenhard
;    APPLICANT:  Morris, Jennifer
;    APPLICANT:  Rauscher III, Frank J.
;    APPLICANT:  Rodeck, Ulrich
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and Methods of
;    TITLE OF INVENTION:  Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/456,907
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/234,783
;      FILING DATE:  28-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48AUSA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-456-907-4

  Query Match             100.0%;  Score 51;  DB 1;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db      126 RMFPNAPYL 134


RESULT 10
PCT-US95-05523-4
; Sequence 4, Application PC/TUS9505523
;  GENERAL INFORMATION:
;    APPLICANT:  The Wistar Institute of Anatomy and Biology
;    TITLE OF INVENTION:  WT1 Monoclonal Antibodies and
;    TITLE OF INVENTION:  Methods of Use Therefor
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/05523
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/234,783
;      FILING DATE:  28-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST48PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  429 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US95-05523-4

  Query Match             100.0%;  Score 51;  DB 5;  Length 429;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 RMFPNAPYL 134


RESULT 11
US-08-102-942A-4
; Sequence 4, Application US/08102942A
; Patent No. 5726288
;  GENERAL INFORMATION:
;    APPLICANT:  Call, Katherine M.
;    APPLICANT:  Glaser, Thomas M.
;    APPLICANT:  Ito, Caryn Y.
;    APPLICANT:  Buckler, Alan J.
;    APPLICANT:  Pelletier, Jerry
;    APPLICANT:  Haber, Daniel A.
;    APPLICANT:  Rose, Elise A.
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Bruening, Wendy
;    APPLICANT:  Darveau, Andre
;    TITLE OF INVENTION:  Localization and Characterization of the
;    TITLE OF INVENTION:  Wilms' Tumor Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,942A
;      FILING DATE:  02-AUG-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5194A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-102-942A-4

  Query Match             100.0%;  Score 51;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 RMFPNAPYL 134


RESULT 12
US-08-102-942A-6
; Sequence 6, Application US/08102942A
; Patent No. 5726288
;  GENERAL INFORMATION:
;    APPLICANT:  Call, Katherine M.
;    APPLICANT:  Glaser, Thomas M.
;    APPLICANT:  Ito, Caryn Y.
;    APPLICANT:  Buckler, Alan J.
;    APPLICANT:  Pelletier, Jerry
;    APPLICANT:  Haber, Daniel A.
;    APPLICANT:  Rose, Elise A.
;    APPLICANT:  Housman, David E.
;    APPLICANT:  Bruening, Wendy
;    APPLICANT:  Darveau, Andre
;    TITLE OF INVENTION:  Localization and Characterization of the
;    TITLE OF INVENTION:  Wilms' Tumor Gene
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
;      STATE:  Massachusetts
;      COUNTRY:  U.S.
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/102,942A
;      FILING DATE:  02-AUG-1993
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-5194A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  449 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-102-942A-6

  Query Match             100.0%;  Score 51;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 RMFPNAPYL 134


RESULT 13
US-09-037-179B-4
; Sequence 4, Application US/09037179B
; Patent No. 6316599
; GENERAL INFORMATION:
;  APPLICANT: Call, Katherine M.
;  APPLICANT:  Glaser, Thomas M.
;  APPLICANT:  Ito, Caryn Y.
;  APPLICANT:  Buckler, Alan J.
;  APPLICANT:  Pelletier, Jerry
;  APPLICANT:  Haber, Daniel A.
;  APPLICANT:  Rose, Elise A.
;  APPLICANT:  Housman, David E.
;  APPLICANT:  Bruening, Wendy
;  APPLICANT:  Darveau, Andre
;  TITLE OF INVENTION: Localization and Characterization of the
;  TITLE OF INVENTION:  Wilms' Tumor Gene
;  FILE REFERENCE: 0050.1312-011
;  CURRENT APPLICATION NUMBER: US/09/037,179B
;  CURRENT FILING DATE:  1998-03-09
;  PRIOR APPLICATION NUMBER: US 08/102,942
;  PRIOR FILING DATE: 1993-08-02
;  PRIOR APPLICATION NUMBER: US 07/614,161
;  PRIOR FILING DATE: 1990-11-13
;  PRIOR APPLICATION NUMBER: US 07/435,780
;  PRIOR FILING DATE: 1989-11-13
;  PRIOR APPLICATION NUMBER: US 07/795,323
;  PRIOR FILING DATE: 1994-09-27
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 449
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Murine
US-09-037-179B-4

  Query Match             100.0%;  Score 51;  DB 4;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 RMFPNAPYL 134


RESULT 14
US-09-037-179B-6
; Sequence 6, Application US/09037179B
; Patent No. 6316599
; GENERAL INFORMATION:
;  APPLICANT: Call, Katherine M.
;  APPLICANT:  Glaser, Thomas M.
;  APPLICANT:  Ito, Caryn Y.
;  APPLICANT:  Buckler, Alan J.
;  APPLICANT:  Pelletier, Jerry
;  APPLICANT:  Haber, Daniel A.
;  APPLICANT:  Rose, Elise A.
;  APPLICANT:  Housman, David E.
;  APPLICANT:  Bruening, Wendy
;  APPLICANT:  Darveau, Andre
;  TITLE OF INVENTION: Localization and Characterization of the
;  TITLE OF INVENTION:  Wilms' Tumor Gene
;  FILE REFERENCE: 0050.1312-011
;  CURRENT APPLICATION NUMBER: US/09/037,179B
;  CURRENT FILING DATE:  1998-03-09
;  PRIOR APPLICATION NUMBER: US 08/102,942
;  PRIOR FILING DATE: 1993-08-02
;  PRIOR APPLICATION NUMBER: US 07/614,161
;  PRIOR FILING DATE: 1990-11-13
;  PRIOR APPLICATION NUMBER: US 07/435,780
;  PRIOR FILING DATE: 1989-11-13
;  PRIOR APPLICATION NUMBER: US 07/795,323
;  PRIOR FILING DATE: 1994-09-27
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 449
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-037-179B-6

  Query Match             100.0%;  Score 51;  DB 4;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 RMFPNAPYL 134


RESULT 15
US-08-975-080-26
; Sequence 26, Application US/08975080
; Patent No. 6245523
;  GENERAL INFORMATION:
;    APPLICANT:  Altieri, Dario C.
;    TITLE OF INVENTION:  SURVIVIN, A PROTEIN THAT INHIBITS
;    TITLE OF INVENTION:  CELLULAR APOPTOSIS, AND ITS MODULATION
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN, LEWIS & BOCKIUS LLP
;      STREET:  1800 M Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036-5869
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/975,080
;      FILING DATE:  20-NOV-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/031,435
;      FILING DATE:  20-NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Adler, Reid G.
;      REGISTRATION NUMBER:  30,988
;      REFERENCE/DOCKET NUMBER:  044574-5022-01-WO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-467-7000
;      TELEFAX:  202-467-7176
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-975-080-26

  Query Match             72.5%;  Score 37;  DB 4;  Length 50;
  Best Local Similarity   66.7%;  Pred. No. 3.7;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
```

```
           | ||| |:|
Db      32 RCFPNCPFL 40

Search completed: April 25, 2003, 06:51:36
Job time : 31 secs
```

```
                  GenCore version 5.1.4_p5_4578
           Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         April 25, 2003, 06:48:04 ; Search time 270 Seconds
                                           (without alignments)
                                           2.671 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters:      301932

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
                1:  /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
                2:  /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
                3:  /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
                4:  /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
                5:  /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
                6:  /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
                7:  /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
                8:  /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
                9:  /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
                10:  /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
                11:  /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
                12:  /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
                13:  /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
                14:  /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                %
Result          Query
  No.   Score   Match Length DB  ID                      Description
--------------------------------------------------------------------------
     1     51   100.0      9  9  US-10-125-635A-185         Sequence 185, App
     2     51   100.0      9  9  US-10-125-635A-293         Sequence 293, App
     3     51   100.0      9  9  US-09-938-864-185          Sequence 185, App
     4     51   100.0      9  9  US-09-938-864-293          Sequence 293, App
     5     51   100.0      9 10   US-09-872-832-46           Sequence 46, Appl
     6     51   100.0     23  9  US-10-125-635A-2           Sequence 2, Appli
     7     51   100.0     23  9  US-10-125-635A-3           Sequence 3, Appli
     8     51   100.0     23  9  US-09-938-864-2            Sequence 2, Appli
     9     51   100.0     23  9  US-09-938-864-3            Sequence 3, Appli
    10     51   100.0    152  9  US-10-125-635A-343         Sequence 343, App
    11     51   100.0    152  9  US-09-938-864-343          Sequence 343, App
    12     51   100.0    154 10   US-09-929-315-5            Sequence 5, Appli
    13     51   100.0    154 10   US-09-993-215-5            Sequence 5, Appli
    14     51   100.0    214  9  US-10-125-635A-395         Sequence 395, App
    15     51   100.0    214  9  US-09-938-864-395          Sequence 395, App
    16     51   100.0    256  9  US-10-125-635A-335         Sequence 335, App
    17     51   100.0    256  9  US-09-938-864-335          Sequence 335; App.
    18     51   100.0    280  9  US-10-125-635A-461         Sequence 461, App
    19     51   100.0    321  9  US-10-125-635A-455         Sequence 455, App
```

```
    20     51   100.0    344  9  US-10-125-635A-391         Sequence 391, App
    21     51   100.0    344  9  US-09-938-864-391          Sequence 391, App
    22     51   100.0    345 10   US-09-929-315-2            Sequence 2, Appli
    23     51   100.0    345 10   US-09-993-215-2            Sequence 2, Appli
    24     51   100.0    362  9  US-10-125-635A-394         Sequence 394, App
    25     51   100.0    362  9  US-09-938-864-394          Sequence 394, App
    26     51   100.0    369  9  US-10-125-635A-346         Sequence 346, App
    27     51   100.0    369  9  US-09-938-864-346          Sequence 346, App
    28     51   100.0    410  9  US-10-125-635A-333         Sequence 333, App
    29     51   100.0    410  9  US-09-938-864-333          Sequence 333, App
    30     51   100.0    420  9  US-10-125-635A-393         Sequence 393, App
    31     51   100.0    420  9  US-09-938-864-393          Sequence 393, App
    32     51   100.0    428  9  US-10-125-635A-405         Sequence 405, App
    33     51   100.0    428  9  US-09-938-864-405          Sequence 405, App
    34     51   100.0    429  9  US-10-125-635A-408         Sequence 408, App
    35     51   100.0    429  9  US-09-938-864-408          Sequence 408, App
    36     51   100.0    449  9  US-10-125-635A-319         Sequence 319, App
    37     51   100.0    449  9  US-10-125-635A-320         Sequence 320, App
    38     51   100.0    449  9  US-10-125-635A-404         Sequence 404, App
    39     51   100.0    449  9  US-09-938-864-319          Sequence 319, App
    40     51   100.0    449  9  US-09-938-864-320          Sequence 320, App
    41     51   100.0    449  9  US-09-938-864-404          Sequence 404, App
    42     51   100.0    449 10   US-09-929-315-4            Sequence 4, Appli
    43     51   100.0    449 10   US-09-929-315-6            Sequence 6, Appli
    44     51   100.0    449 10   US-09-993-215-4            Sequence 4, Appli
    45     51   100.0    449 10   US-09-993-215-6            Sequence 6, Appli

                                  ALIGNMENTS


RESULT 1
US-10-125-635A-185
; Sequence 185, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-125-635A-185

  Query Match             100.0%;  Score 51;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 RMFPNAPYL 9
             |||||||||
Db         1 RMFPNAPYL 9


RESULT 2
US-10-125-635A-293
; Sequence 293, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
```

```
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-125-635A-293

  Query Match            100.0%;  Score 51;  DB 9;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9


RESULT 3
US-09-938-864-185
; Sequence 185, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 185
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-185

  Query Match            100.0%;  Score 51;  DB 9;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9


RESULT 4
US-09-938-864-293
; Sequence 293, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 293
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-293

  Query Match            100.0%;  Score 51;  DB 9;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9


RESULT 5
US-09-872-832-46
; Sequence 46, Application US/09872832
; Patent No. US20020131960A1
; GENERAL INFORMATION:
;  APPLICANT: Memorial Sloan-Kettering Cancer Center
;  TITLE OF INVENTION: ARTIFICIAL ANTIGEN PRESENTING CELLS AND METHODS OF USE THEREOF
;  FILE REFERENCE: 830002-2003.1
;  CURRENT APPLICATION NUMBER: US/09/872,832
;  CURRENT FILING DATE:  2001-06-01
;  PRIOR APPLICATION NUMBER: 60/209,157
;  PRIOR FILING DATE: 2000-02-06
;  NUMBER OF SEQ ID NOS: 49
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 46
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-872-832-46

  Query Match            100.0%;  Score 51;  DB 10;  Length 9;
  Best Local Similarity  100.0%;  Pred. No. 2.7e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db        1 RMFPNAPYL 9


RESULT 6
US-10-125-635A-2
; Sequence 2, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
```

```
;   LENGTH: 23
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-125-635A-2

  Query Match             100.0%;  Score 51;  DB 9;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.0087;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      10 RMFPNAPYL 18

RESULT 7
US-10-125-635A-3
; Sequence 3, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 23
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-125-635A-3

  Query Match             100.0%;  Score 51;  DB 9;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.0087;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      10 RMFPNAPYL 18

RESULT 8
US-09-938-864-2
; Sequence 2, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 23
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-938-864-2

  Query Match             100.0%;  Score 51;  DB 9;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.0087;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      10 RMFPNAPYL 18

RESULT 9
US-09-938-864-3
; Sequence 3, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 23
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-938-864-3

  Query Match             100.0%;  Score 51;  DB 9;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.0087;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      10 RMFPNAPYL 18

RESULT 10
US-10-125-635A-343
; Sequence 343, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 343
;   LENGTH: 152
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-125-635A-343

  Query Match             100.0%;  Score 51;  DB 9;  Length 152;
```

```
  Best Local Similarity   100.0%;  Pred. No. 0.059;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      51 RMFPNAPYL 59


RESULT 11
US-09-938-864-343
; Sequence 343, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 343
;   LENGTH: 152
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-343

  Query Match             100.0%;  Score 51;  DB 9;  Length 152;
  Best Local Similarity   100.0%;  Pred. No. 0.059;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      51 RMFPNAPYL 59


RESULT 12
US-09-929-315-5
; Sequence 5, Application US/09929315
; Patent No. US20020082394A1
; GENERAL INFORMATION:
;  APPLICANT: Call, Katherine M.
;  APPLICANT:  Glaser, Thomas M.
;  APPLICANT:  Ito, Caryn Y.
;  APPLICANT:  Buckler, Alan J.
;  APPLICANT:  Pelletier, Jerry
;  APPLICANT:  Haber, Daniel A.
;  APPLICANT:  Rose, Elise A.
;  APPLICANT:  Housman, David E.
;  APPLICANT:  Bruening, Wendy
;  APPLICANT:  Darveau, Andre
;  TITLE OF INVENTION: Localization and Characterization of the
;  TITLE OF INVENTION:  Wilms' Tumor Gene
;  FILE REFERENCE: 0050.1312-013
;  CURRENT APPLICATION NUMBER: US/09/929,315
;  CURRENT FILING DATE:  2001-08-14
;  PRIOR APPLICATION NUMBER: US 09/037,179
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 08/102,942
;  PRIOR FILING DATE: 1993-08-02
;  PRIOR APPLICATION NUMBER: US 07/614,161
;  PRIOR FILING DATE: 1990-11-13
;  PRIOR APPLICATION NUMBER: US 07/435,780
;  PRIOR FILING DATE: 1989-11-13
;  PRIOR APPLICATION NUMBER: US 07/795,323
;  PRIOR FILING DATE: 1991-11-20
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 154
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-929-315-5

  Query Match             100.0%;  Score 51;  DB 10;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 0.059;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      40 RMFPNAPYL 48


RESULT 13
US-09-993-215-5
; Sequence 5, Application US/09993215
; Patent No. US20020128196A1
; GENERAL INFORMATION:
;  APPLICANT: Call, Katherine M.
;  APPLICANT:  Glaser, Thomas M.
;  APPLICANT:  Ito, Caryn Y.
;  APPLICANT:  Buckler, Alan J.
;  APPLICANT:  Pelletier, Jerry
;  APPLICANT:  Haber, Daniel A.
;  APPLICANT:  Rose, Elise A.
;  APPLICANT:  Housman, David E.
;  APPLICANT:  Bruening, Wendy
;  APPLICANT:  Darveau, Andre
;  TITLE OF INVENTION: Localization and Characterization of the
;  TITLE OF INVENTION:  Wilms' Tumor Gene
;  FILE REFERENCE: 0050.1312-014
;  CURRENT APPLICATION NUMBER: US/09/993,215
;  CURRENT FILING DATE:  2001-11-12
;  PRIOR APPLICATION NUMBER: US 09/037,179
;  PRIOR FILING DATE: 1998-03-09
;  PRIOR APPLICATION NUMBER: US 08/102,942
;  PRIOR FILING DATE: 1993-08-02
;  PRIOR APPLICATION NUMBER: US 07/614,161
;  PRIOR FILING DATE: 1990-11-13
;  PRIOR APPLICATION NUMBER: US 07/435,780
;  PRIOR FILING DATE: 1989-11-13
;  PRIOR APPLICATION NUMBER: US 07/795,323
;  PRIOR FILING DATE: 1991-11-20
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 154
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-993-215-5

  Query Match             100.0%;  Score 51;  DB 10;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 0.059;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db      40 RMFPNAPYL 48


RESULT 14
US-10-125-635A-395
; Sequence 395, Application US/10125635A
; Publication No. US20030039635A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  Smithgall, Molly D.
```

```
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Cheever, Martin A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Sutherland, R. Alec
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C7
;  CURRENT APPLICATION NUMBER: US/10/125,635A
;  CURRENT FILING DATE:  2002-07-19
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 395
;   LENGTH: 214
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-125-635A-395

  Query Match             100.0%;  Score 51;  DB 9;  Length 214;
  Best Local Similarity   100.0%;  Pred. No. 0.083;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db     59 RMFPNAPYL 67


RESULT 15
US-09-938-864-395
; Sequence 395, Application US/09938864
; Publication No. US20030072767A1
; GENERAL INFORMATION:
;  APPLICANT: Gaiger, Alexander
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Smithgall, Molly
;  APPLICANT:  Moulton, Gus
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Sleath, Paul R.
;  APPLICANT:  Mossman, Sally
;  APPLICANT:  Evans, Lawrence
;  APPLICANT:  Spies, A. Gregory
;  APPLICANT:  Boydston, Jeremy
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR WT1
;  TITLE OF INVENTION:  SPECIFIC IMMUNOTHERAPY
;  FILE REFERENCE: 210121.465C5
;  CURRENT APPLICATION NUMBER: US/09/938,864
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 413
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 395
;   LENGTH: 214
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-938-864-395

  Query Match             100.0%;  Score 51;  DB 9;  Length 214;
  Best Local Similarity   100.0%;  Pred. No. 0.083;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db     59 RMFPNAPYL 67

Search completed: April 25, 2003, 06:57:02
Job time : 270 secs
```

```
                    GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         April 25, 2003, 06:36:22 ; Search time 75 Seconds
                                           (without alignments)
                                           15.990 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:      908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_101002:*
                 1:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                 2:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                 3:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                 4:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                 5:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                 6:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                 7:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                 8:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                 9:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                 10:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                 11:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                 12:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                 13:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                 14:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                 15:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                 16:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                 17:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                 18:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                 19:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                 20:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                 21:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                 22:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                 23:  /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES
                 %
Result          Query
  No.   Score   Match Length DB  ID                     Description
-----------------------------------------------------------------------
    1      51   100.0      9  21  AAY94202               Human cytotoxic T
    2      51   100.0      9  21  AAY98670               WT1 derived immuno
    3      51   100.0      9  21  AAY98778               WT1 derived immuno
    4      51   100.0      9  21  AAY80200               Human Wilms' tumou
    5      51   100.0      9  22  AAU68769               Human Wilm's tumou
    6      51   100.0      9  22  AAU68877               Mouse Wilm's tumou
    7      51   100.0      9  22  AAG62002               Human WT1 immunoge
    8      51   100.0      9  22  AAG62110               Mouse WT1 immunoge
    9      51   100.0      9  23  ABG33239               Human WT1 immunoge
   10      51   100.0      9  23  ABG33347               Mouse WT1 immunoge
   11      51   100.0      9  23  AAE17298               Human leukocyte an
   12      51   100.0     23  21  AAY98502               Human WT1 peptide
   13      51   100.0     23  21  AAY98503               Mouse WT1 peptide
   14      51   100.0     23  22  AAU68601               Human Wilm's tumou
   15      51   100.0     23  22  AAU68602               Mouse Wilm's tumou
   16      51   100.0     23  22  AAG61834               Human WT1 immunoge
   17      51   100.0     23  22  AAG61835               Mouse WT1 immunoge
   18      51   100.0     23  23  ABG33071               Human WT1 immunoge
   19      51   100.0     23  23  ABG33072               Mouse WT1 immunoge
   20      51   100.0    152  23  ABG33387               Human WT1-Tr2.  Ho
   21      51   100.0    154  19  AAW47175               Wilms' tumour poly
   22      51   100.0    154  23  AAG78445               Amino acid sequenc
   23      51   100.0    210  16  AAR85066               Wilms' tumour anti
   24      51   100.0    210  18  AAW22883               Wilms' tumour prot
   25      51   100.0    214  23  ABG33395               Human WT1-E.  Homo
   26      51   100.0    256  23  ABG33384               Human WT1-A.  Homo
   27      51   100.0    344  23  ABG33391               Human Ra12/WT1-E.
   28      51   100.0    345  12  AAR12240               Wilm's tumour gene
   29      51   100.0    345  19  AAW47173               Wilms' tumour poly
   30      51   100.0    345  23  AAG78443               WT33 Wilm's tumour
   31      51   100.0    362  23  ABG33394               Human WT1-C.  Homo
   32      51   100.0    369  23  ABG33390               Human WT1-C.  Homo
   33      51   100.0    410  23  ABG33382               Human Trx_WT1_A.
   34      51   100.0    420  23  ABG33393               Human Ra12/WT1-F.
   35      51   100.0    428  23  ABG33397               Human WT1 variant
   36      51   100.0    429  16  AAR85065               Wilms' tumour WT1
   37      51   100.0    429  23  ABG33400               Human WT1 protein,
   38      51   100.0    449  19  AAW47176               Wilms' tumour poly
   39      51   100.0    449  21  AAY98804               Human WT1 protein
   40      51   100.0    449  21  AAY98805               Mouse WT1 protein
   41      51   100.0    449  21  AAY80196               Mouse Wilms' tumou
   42      51   100.0    449  21  AAY80197               Human Wilms' tumou
   43      51   100.0    449  22  AAU68903               Human Wilm's tumou
   44      51   100.0    449  22  AAU68904               Mouse Wilm's tumou
   45      51   100.0    449  22  AAG62136               Human WT1 protein

                                 ALIGNMENTS


RESULT 1
AAY94202
ID   AAY94202 standard; peptide; 9 AA.
XX
AC   AAY94202;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human cytotoxic T lymphocyte-recognised WT1 peptide WT126-34.
XX
KW   WT126-3;.peptide; epitope; Wilm's tumour gene; leukaemia;
KW   breast cancer; melanoma; ovarian cancer; immunotherapy.
XX
OS   Homo sapiens.
XX
PN   WO200026249-A1.
XX
PD   11-MAY-2000.
XX
PF   02-NOV-1999;   99WO-GB03572.
XX
PR   02-NOV-1998;   98GB-0023897.
XX
PA   (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI   Stauss HJ,  Gao L;
XX
DR   WPI; 2000-376123/32.
XX
PT   Novel peptides comprising WT-1 and GATA-1 epitopes, their fragments or
PT   variants, useful as vaccines for cancer immunotherapy  -
XX
PS   Claim 1; Page 74; 93pp; English.
```

```
XX
CC   The present sequence is peptide epitope WT126-34, produced by WT1
CC   expressing cells and found at residues 126-134 of the WT1 protein, which
CC   is recognised by cytotoxic T lymphocytes. WT1 is aberrantly expressed in
CC   leukaemias, breast cancer, melanoma and ovarian cancer. The peptide can
CC   be used as a vaccine to stimulate the elimination, by cytotoxic T
CC   lymphocytes, of cancer cells aberrantly expressing WT1. In addition, the
CC   nucleic acid encoding the peptide may also be used in the same manner.
CC   Alternatively, the peptide may be used in vitro to produce activated
CC   cytotoxic T lymphocytes.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 RMFPNAPYL 9
             |||||||||
Db         1 RMFPNAPYL 9


RESULT 2
AAY98670
ID   AAY98670 standard; Peptide; 9 AA.
XX
AC   AAY98670;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:185.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 171; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
```

```
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 RMFPNAPYL 9
             |||||||||
Db         1 RMFPNAPYL 9


RESULT 3
AAY98778
ID   AAY98778 standard; Peptide; 9 AA.
XX
AC   AAY98778;
XX
DT   31-JUL-2000  (first entry)
XX
DE   WT1 derived immunogenic peptide SEQ ID NO:293.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 186; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
```

```
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             |||||||||
Db       1 RMFPNAPYL 9


RESULT 4
AAY80200
ID   AAY80200 standard; Peptide; 9 AA.
XX
AC   AAY80200;
XX
DT   24-MAY-2000  (first entry)
XX
DE   Human Wilms' tumour suppressor gene WT1 product peptide SEQ ID NO:5.
XX
KW   Wilm's tumour suppressor gene; WT1; cancer; antigen; vaccine; MHC;
KW   major histocompatibility complex; leukaemia; tumour; antitumour.
XX
OS   Homo sapiens.
XX
PN   WO200006602-A1.
XX
PD   10-FEB-2000.
XX
PF   30-JUL-1999;   99WO-JP04130.
XX
PR   31-JUL-1998;   98JP-0218093.
XX
PA   (SUGI/) SUGIYAMA H.
XX
PI   Sugiyama H,  Oka Y;
XX
DR   WPI; 2000-195264/17.
XX
PT   Cancer antigens based on Wilm's tumor suppressor gene WT1 product or
PT   peptide derivatives, for cancer vaccines in treating leukemia and solid
PT   tumors e.g. stomach cancer, skin cancer, lung cancer and breast cancer
XX
PS   Claim 5; Page 18; 48pp; Japanese.
XX
CC   The present invention describes a cancer antigen containing the active
CC   component of Wilms' tumour suppressor gene WT1 product, or partial
CC   peptides, for cancer vaccines in treating leukaemia and solid tumours.
CC   The cancer antigens are useful for cancer vaccines in treating
CC   leukaemia, bone-marrow abnormal formation syndrome, malignant lymphoma,
CC   multiple myeloma, stomach cancer, cancer of the large intestine, lung
CC   cancer, breast cancer, blastoma, liver cancer, skin cancer, bladder
CC   cancer, prostate cancer, uterus cancer, cervical cancer, or ovary
CC   cancer. The present sequence represents a peptide from the human
CC   Wilms' tumour suppressor gene WT1 product.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             |||||||||
Db       1 RMFPNAPYL 9
```

```
RESULT 5
AAU68769
ID   AAU68769 standard; Peptide; 9 AA.
XX
AC   AAU68769;
XX
DT   16-JAN-2002  (first entry)
XX
DE   Human Wilm's tumour protein, WT1, antigenic peptide #164.
XX
KW   Human; Wilms' tumour; WT1; pleural mesothelioma; antigen;
KW   leukaemia; acute myeloid leukaemia; AML; chronic myeloid leukaemia; CML;
KW   acute lymphocytic leukaemia; ALL; myelodysplastic syndromes;
KW   myeloproliferative syndrome; cancer; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200162920-A2.
XX
PD   30-AUG-2001.
XX
PF   22-FEB-2001; 2001WO-US05702.
XX
PR   22-FEB-2000; 2000US-184070P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Cheever MA,  Gaiger A;
XX
DR   WPI; 2001-648218/74.
XX
PT   Composition for the treatment of mesothelioma comprises specific
PT   peptides i.e. Wilms' tumour antigen polypeptide derived antigenic
PT   fragments -
XX
PS   Claim 1; Page 176; 242pp; English.
XX
CC   The invention relates to the use of a composition comprising at least a
CC   first isolated peptide, of between 9 and 40 amino acids or a first
CC   nucleic acid, encoding the peptide, in the manufacture of a medicament
CC   for treating or preventing mesothelioma. The peptides are antigenic
CC   peptides derived from the Wilms' tumour protein WT1. The composition is
CC   useful for the treatment of mesothelioma, Wilms' tumour, preferably
CC   pleural mesothelioma and other WT1 associated malignancies e.g.
CC   leukaemia (including acute myeloid leukaemia, AML, chronic myeloid
CC   leukaemia, CML, acute lymphocytic leukaemia, ALL, and childhood ALL),
CC   myelodysplastic syndromes, myeloproliferative syndromes and cancers
CC   (e.g. breast, testicular, prostate, lung and ovarian) in mammals,
CC   preferably humans. The present sequence is an antigenic peptide of
CC   the invention derived from human WT1.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             |||||||||
Db       1 RMFPNAPYL 9


RESULT 6
AAU68877
ID   AAU68877 standard; Peptide; 9 AA.
XX
AC   AAU68877;
XX
DT   16-JAN-2002  (first entry)
XX
DE   Mouse Wilm's tumour protein, WT1, antigenic peptide #46.
XX
KW   Mouse; Wilms' tumour; WT1; pleural mesothelioma; antigen;
```

```
KW   leukaemia; acute myeloid leukaemia; AML; chronic myeloid leukaemia; CML;
KW   acute lymphocytic leukaemia; ALL; myelodysplastic syndromes;
KW   myeloproliferative syndrome; cancer; cytostatic.
XX
OS   Mus musculus.
XX
PN   WO200162920-A2.
XX
PD   30-AUG-2001.
XX
PF   22-FEB-2001; 2001WO-US05702.
XX
PR   22-FEB-2000; 2000US-184070P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Cheever MA,  Gaiger A;
XX
DR   WPI; 2001-648218/74.
XX
PT   Composition for the treatment of mesothelioma comprises specific
PT   peptides i.e. Wilms' tumour antigen polypeptide derived antigenic
PT   fragments -
XX
PS   Claim 1; Page 24; 242pp; English.
XX
CC   The invention relates to the use of a composition comprising at least a
CC   first isolated peptide, of between 9 and 40 amino acids or a first
CC   nucleic acid, encoding the peptide, in the manufacture of a medicament
CC   for treating or preventing mesothelioma. The peptides are antigenic
CC   peptides derived from the Wilms' tumour protein WT1. The composition is
CC   useful for the treatment of mesothelioma, Wilms' tumour, preferably
CC   pleural mesothelioma and other WT1 associated malignancies e.g.
CC   leukaemia (including acute myeloid leukaemia, AML, chronic myeloid
CC   leukaemia, CML, acute lymphocytic leukaemia, ALL, and childhood ALL),
CC   myelodysplastic syndromes, myeloproliferative syndromes and cancers
CC   (e.g. breast, testicular, prostate, lung and ovarian) in mammals,
CC   preferably humans. The present sequence is an antigenic peptide of
CC   the invention derived from mouse WT1.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db          1 RMFPNAPYL 9


RESULT 7
AAG62002
ID   AAG62002 standard; Peptide; 9 AA.
XX
AC   AAG62002;
XX
DT   06-JUL-2001  (first entry)
XX
DE   Human WT1 immunogenic peptide SEQ ID NO: 185.
XX
KW   Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;
KW   chromosome 11p13; zinc finger transcription factor.
XX
OS   Homo sapiens.
XX
PN   WO200125273-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27465.
XX
PR   04-OCT-1999;    99US-0157459.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Xu J,  Cheever MA,  Reed SG;
XX
DR   WPI; 2001-328324/34.
XX
PT   Polypeptide comprising part of the Wilm's Tumour gene product sequence is
PT   used in the diagnosis and treatment of malignant diseases e.g. leukemia
PT   and cancer associated with WTI  -
XX
PS   Claim 4; Page 183; 228pp; English.
XX
CC   The present invention describes compositions comprising peptides derived
CC   from the Wilm's tumour protein WT1 and methods for their use in treating
CC   malignant diseases. Peptides derived from both the murine and human WT1
CC   proteins are provided. The human WT1 gene is found on chromosome 11p13,
CC   and the protein was shown to be a zinc finger transcription factor. The
CC   immunogenic peptides of the invention are particularly useful in the
CC   diagnosis and treatment of cancer and leukaemia. The present sequence is
CC   a polypeptide described in the exemplification of the invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db          1 RMFPNAPYL 9


RESULT 8
AAG62110
ID   AAG62110 standard; Peptide; 9 AA.
XX
AC   AAG62110;
XX
DT   06-JUL-2001  (first entry)
XX
DE   Mouse WT1 immunogenic peptide SEQ ID NO: 293.
XX
KW   Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilm's tumour gene;
KW   chromosome 11p13; zinc finger transcription factor.
XX
OS   Mus musculus.
XX
PN   WO200125273-A2.
XX
PD   12-APR-2001.
XX
PF   04-OCT-2000; 2000WO-US27465.
XX
PR   04-OCT-1999;    99US-0157459.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Skeiky YAW,  Xu J,  Cheever MA,  Reed SG;
XX
DR   WPI; 2001-328324/34.
XX
PT   Polypeptide comprising part of the Wilm's Tumour gene product sequence is
PT   used in the diagnosis and treatment of malignant diseases e.g. leukemia
PT   and cancer associated with WTI  -
XX
PS   Claim 4; Page 200; 228pp; English.
XX
CC   The present invention describes compositions comprising peptides derived
CC   from the Wilm's tumour protein WT1 and methods for their use in treating
CC   malignant diseases. Peptides derived from both the murine and human WT1
CC   proteins are provided. The human WT1 gene is found on chromosome 11p13,
CC   and the protein was shown to be a zinc finger transcription factor. The
```

```
CC   immunogenic peptides of the invention are particularly useful in the
CC   diagnosis and treatment of cancer and leukaemia. The present sequence is
CC   a polypeptide described in the exemplification of the invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
                |||||||||
Db          1 RMFPNAPYL 9


RESULT 9
ABG33239
ID   ABG33239 standard; Peptide; 9 AA.
XX
AC   ABG33239;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human WT1 immunogenic peptide #162.
XX
KW   Human; mouse; cytostatic; immunostimulant;  WT1; cancer;
KW   immune response.
XX
OS   Homo sapiens.
XX
PN   WO200228414-A1.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US31139.
XX
PR   06-OCT-2000; 2000US-0684361.
PR   09-OCT-2000; 2000US-0685830.
PR   15-FEB-2001; 2001US-0785019.
PR   24-AUG-2001; 2001US-0938864.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  McNeill PD,  Smithgall M,  Moulton G,  Vedvick TS;
PI   Sleath PR,  Mossman S,  Evans L,  Spies AG,  Boydston J;
XX
DR   WPI; 2002-352217/38.
XX
PT   Novel isolated WT1 polynucleotide, and encoded polypeptide, useful for
PT   treating and diagnosing cancer in a patient -
XX
PS   Example 4; Page 194; 260pp; English.
XX
CC   The invention relates to an isolated WT1 polynucleotide (I) and
CC   polypeptide encoded by (I). The WT1 polynucleotides and polypeptides
CC   are used for treating and detecting cancer in a patient, and for
CC   stimulating an immune response in patient. ABG33070-ABG33405
CC   represent WT1 amino acid sequences of the invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 23;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
                |||||||||
Db          1 RMFPNAPYL 9


RESULT 10
ABG33347
ID   ABG33347 standard; Peptide; 9 AA.
XX
AC   ABG33347;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Mouse WT1 immunogenic peptide #44.
XX
KW   Human; mouse; cytostatic; immunostimulant;  WT1; cancer;
KW   immune response.
XX
OS   Mus musculus.
XX
PN   WO200228414-A1.
XX
PD   11-APR-2002.
XX
PF   03-OCT-2001; 2001WO-US31139.
XX
PR   06-OCT-2000; 2000US-0684361.
PR   09-OCT-2000; 2000US-0685830.
PR   15-FEB-2001; 2001US-0785019.
PR   24-AUG-2001; 2001US-0938864.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  McNeill PD,  Smithgall M,  Moulton G,  Vedvick TS;
PI   Sleath PR,  Mossman S,  Evans L,  Spies AG,  Boydston J;
XX
DR   WPI; 2002-352217/38.
XX
PT   Novel isolated WT1 polynucleotide, and encoded polypeptide, useful for
PT   treating and diagnosing cancer in a patient -
XX
PS   Example 4; Page 210; 260pp; English.
XX
CC   The invention relates to an isolated WT1 polynucleotide (I) and
CC   polypeptide encoded by (I). The WT1 polynucleotides and polypeptides
CC   are used for treating and detecting cancer in a patient, and for
CC   stimulating an immune response in patient. ABG33070-ABG33405
CC   represent WT1 amino acid sequences of the invention.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 23;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
                |||||||||
Db          1 RMFPNAPYL 9


RESULT 11
AAE17298
ID   AAE17298 standard; peptide; 9 AA.
XX
AC   AAE17298;
XX
DT   18-APR-2002  (first entry)
XX
DE   Human leukocyte antigen (HLA-A2.1) restricted peptide, Db126.
XX
KW   Human; artificial antigen presenting cell; AAPC; beta2-microglobulin;
KW   human leukocyte antigen; HLA; major histocompatibility complex; MHC;
KW   cytotoxic T lymphocyte; CTL; T cell-specific antigen; TCA; antitumour;
KW   immune response; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200194944-A2.
XX
```

```
PD   13-DEC-2001.
XX
PF   01-JUN-2001; 2001WO-US17981.
XX
PR   02-JUN-2000; 2000US-209157P.
XX
PA   (SLOK ) SLOAN KETTERING INST CANCER RES.
XX
PI   Sadelain M,  Latouche J;
XX
DR   WPI; 2002-139667/18.
XX
PT   Artificial antigen presenting cells for activating T lymphocytes,
PT   comprises eukaryotic cell expressing antigen presenting complex having
PT   beta2-microglobulin, exogenous accessory molecule, human leukocyte
PT   antigen molecule and protein -
XX
PS   Example 17; Page 40; 75pp; English.
XX
CC   The present invention relates to an artificial antigen presenting cell
CC   (AAPC) comprising a eukaryotic cell expressing an antigen presenting
CC   complex comprising beta2-microglobulin, an exogenous accessory molecule,
CC   a human leukocyte antigen, HLA (major histocompatibility complex, MHC)
CC   molecule of a single type and a protein that is processed intracellularly
CC   to produce an exogenous T cell-specific epitope. The invention also
CC   relates to methods for activation of T lymphocytes. The method is also
CC   useful for identifying within a test population of cytotoxic T
CC   lymphocytes (CTLs), CTLs specifically activated against a known T-cell
CC   specific antigen (TCA), which is useful for diagnostic purposes. AAPC is
CC   also useful for activating CTLs, by contacting AAPC with a suitable
CC   population of T lymphocytes under conditions suitable for the activation
CC   and isolating the activated CTLs. AAPC is further useful for the
CC   investigation of primary T cell activation and diagnostic applications
CC   here primary T cell activation allow discovery of antigens and accessory
CC   molecules, and diagnostic applications include cell-based assays for
CC   quantifying immune response in normal, infected or treated (vaccinated)
CC   patients. Composition comprising AAPC or activated T cells produced by
CC   utilising AAPC is useful for eliciting an antitumour response. The
CC   invention is used for the treatment of cancer. The present sequence is
CC   human HLA-A2.1 restricted peptide used in additional AAPC-induced CTL-
CC   activation.
XX
SQ   Sequence   9 AA;

  Query Match             100.0%;  Score 51;  DB 23;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 7.8e+05;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db          1 RMFPNAPYL 9


RESULT 12
AAY98502
ID   AAY98502 standard; Peptide; 23 AA.
XX
AC   AAY98502;
XX
DT   31-JUL-2000  (first entry)
XX
DE   Human WT1 peptide SEQ ID NO:2.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
```

```
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Claim 4; Page 46; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   23 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.006;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |||||||||
Db         10 RMFPNAPYL 18


RESULT 13
AAY98503
ID   AAY98503 standard; Peptide; 23 AA.
XX
AC   AAY98503;
XX
DT   31-JUL-2000  (first entry)
XX
DE   Mouse WT1 peptide SEQ ID NO:3.
XX
KW   WT1; immunotherapy; immunogenic; malignant disease; cancer; leukaemia;
KW   metastatic disease; mouse; human; Wilm's tumour; immune response;
KW   vaccine.
XX
OS   Mus musculus.
XX
PN   WO200018795-A2.
XX
PD   06-APR-2000.
XX
PF   30-SEP-1999;   99WO-US22819.
XX
PR   30-SEP-1998;   98US-0164223.
```

```
PR   25-MAR-1999;   99US-0276484.
XX
PA   (CORI-) CORIXA CORP.
PA   (GAIG/) GAIGER A.
XX
PI   Gaiger A,  Cheever M;
XX
DR   WPI; 2000-293107/25.
XX
PT   Novel polypeptides comprising an immunogenic portion of a native WT1
PT   polypeptide, useful for inhibiting the development of malignant
PT   diseases associated with WT1 expression e.g. leukemia or cancer  -
XX
PS   Example 3; Page 46; 193pp; English.
XX
CC   The present invention describes polypeptides (I) comprising an
CC   immunogenic portion of a native Wilm's Tumour gene product polypeptide,
CC   WT1, (or variants of the immunogenic portion retaining the ability to
CC   react with WT1-specific antisera and/or T-cell lines or clones) and
CC   comprising 16 consecutive amino acids (aa) or less of a native WT1
CC   polypeptide. The polypeptides are useful therapeutically and to
CC   manufacture medicaments for enhancing/inducing an immune response in
CC   patients. The polypeptides, mimetics or polynucleotides can be included
CC   with a carrier/excipient in pharmaceutical compositions or with a
CC   non-specific immune response enhancer (e.g. an adjuvant or enhancer
CC   preferentially enhancing a T cell response) in vaccines. Pharmaceutical
CC   compositions and vaccines can be administered to human patients to
CC   enhance or induce an immune response specific for WT1 or a cell
CC   expressing WT1, useful to inhibit the development of malignant diseases
CC   associated with WT1 expression, e.g. leukemia (especially acute/chronic
CC   myeloid leukemia or acute lymphocytic leukemia) or cancer (especially
CC   breast, lung, thyroid or gastrointestinal cancer, or a melanoma).
CC   AAY98501 to AAY98811 represent polypeptide sequences, and AAA13848 to
CC   AAA13862 represent PCR primers, used in the exemplification of the
CC   present invention.
XX
SQ   Sequence   23 AA;

  Query Match             100.0%;  Score 51;  DB 21;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.006;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             |||||||||
Db      10 RMFPNAPYL 18


RESULT 14
AAU68601
ID   AAU68601 standard; Peptide; 23 AA.
XX
AC   AAU68601;
XX
DT   16-JAN-2002  (first entry)
XX
DE   Human Wilm's tumour protein, WT1, antigenic peptide #2.
XX
KW   Human; Wilms' tumour; WT1; pleural mesothelioma; antigen;
KW   leukaemia; acute myeloid leukaemia; AML; chronic myeloid leukaemia; CML;
KW   acute lymphocytic leukaemia; ALL; myelodysplastic syndromes;
KW   myeloproliferative syndrome; cancer; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO200162920-A2.
XX
PD   30-AUG-2001.
XX
PF   22-FEB-2001; 2001WO-US05702.
XX
PR   22-FEB-2000; 2000US-184070P.
XX
PA   (CORI-) CORIXA CORP.
```

```
XX
PI   Cheever MA,  Gaiger A;
XX
DR   WPI; 2001-648218/74.
XX
PT   Composition for the treatment of mesothelioma comprises specific
PT   peptides i.e. Wilms' tumour antigen polypeptide derived antigenic
PT   fragments -
XX
PS   Claim 1; Page 24; 242pp; English.
XX
CC   The invention relates to the use of a composition comprising at least a
CC   first isolated peptide, of between 9 and 40 amino acids or a first
CC   nucleic acid, encoding the peptide, in the manufacture of a medicament
CC   for treating or preventing mesothelioma. The peptides are antigenic
CC   peptides derived from the Wilms' tumour protein WT1. The composition is
CC   useful for the treatment of mesothelioma, Wilms' tumour, preferably
CC   pleural mesothelioma and other WT1 associated malignancies e.g.
CC   leukaemia (including acute myeloid leukaemia, AML, chronic myeloid
CC   leukaemia, CML, acute lymphocytic leukaemia, ALL, and childhood ALL),
CC   myelodysplastic syndromes, myeloproliferative syndromes and cancers
CC   (e.g. breast, testicular, prostate, lung and ovarian) in mammals,
CC   preferably humans. The present sequence is an antigenic peptide of
CC   the invention derived from human WT1.
XX
SQ   Sequence   23 AA;

  Query Match             100.0%;  Score 51;  DB 22;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.006;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             |||||||||
Db      10 RMFPNAPYL 18


RESULT 15
AAU68602
ID   AAU68602 standard; Peptide; 23 AA.
XX
AC   AAU68602;
XX
DT   16-JAN-2002  (first entry)
XX
DE   Mouse Wilm's tumour protein, WT1, antigenic peptide #1.
XX
KW   Mouse; Wilms' tumour; WT1; pleural mesothelioma; antigen;
KW   leukaemia; acute myeloid leukaemia; AML; chronic myeloid leukaemia; CML;
KW   acute lymphocytic leukaemia; ALL; myelodysplastic syndromes;
KW   myeloproliferative syndrome; cancer; cytostatic.
XX
OS   Mus musculus.
XX
PN   WO200162920-A2.
XX
PD   30-AUG-2001.
XX
PF   22-FEB-2001; 2001WO-US05702.
XX
PR   22-FEB-2000; 2000US-184070P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Cheever MA,  Gaiger A;
XX
DR   WPI; 2001-648218/74.
XX
PT   Composition for the treatment of mesothelioma comprises specific
PT   peptides i.e. Wilms' tumour antigen polypeptide derived antigenic
PT   fragments -
XX
PS   Claim 1; Page 24; 242pp; English.
XX
```

```
CC   The invention relates to the use of a composition comprising at least a
CC   first isolated peptide, of between 9 and 40 amino acids or a first
CC   nucleic acid, encoding the peptide, in the manufacture of a medicament
CC   for treating or preventing mesothelioma. The peptides are antigenic
CC   peptides derived from the Wilms' tumour protein WT1. The composition is
CC   useful for the treatment of mesothelioma, Wilms' tumour, preferably
CC   pleural mesothelioma and other WT1 associated malignancies e.g.
CC   leukaemia (including acute myeloid leukaemia, AML, chronic myeloid
CC   leukaemia, CML, acute lymphocytic leukaemia, ALL, and childhood ALL),
CC   myelodysplastic syndromes, myeloproliferative syndromes and cancers
CC   (e.g. breast, testicular, prostate, lung and ovarian) in mammals,
CC   preferably humans. The present sequence is an antigenic peptide of
CC   the invention derived from mouse WT1.
XX
SQ   Sequence   23 AA;

  Query Match            100.0%;  Score 51;  DB 22;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.006;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          |||||||||
Db     10 RMFPNAPYL 18


Search completed: April 25, 2003, 06:47:57
Job time : 77 secs
```

```
                     GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         April 25, 2003, 06:36:22 ; Search time 11 Seconds
                                           (without alignments)
                                           33.935 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters:      112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_40:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES

                 %
Result          Query
  No.   Score  Match Length DB  ID                        Description
-------------------------------------------------------------------------
    1     51   100.0    448  1  WT1_RAT                   P49952 rattus norv
    2     51   100.0    449  1  WT1_HUMAN                 P19544 homo sapien
    3     51   100.0    449  1  WT1_MOUSE                 P22561 mus musculu
    4     51   100.0    449  1  WT1_PIG                   O62651 sus scrofa
    5     40    78.4    262  1  YFIO_HAEIN                P44553 haemophilus
    6     37    72.5    468  1  IFT2_CRILO                Q60462 cricetulus
    7     37    72.5    854  1  UN33_CAEEL                Q01630 caenorhabdi
    8     37    72.5   1403  1  BIR1_HUMAN                Q13075 homo sapien
    9     36    70.6    217  1  GTM1_HUMAN                P09488 homo sapien
   10     36    70.6    217  1  GTM1_MOUSE                P10649 mus musculu
   11     36    70.6    217  1  GTM1_RAT                  P04905 rattus norv
   12     36    70.6    217  1  GTM2_HUMAN                P28161 homo sapien
   13     36    70.6    217  1  GTM2_MOUSE                P15626 mus musculu
   14     36    70.6    217  1  GTM2_RAT                  P08010 rattus norv
   15     36    70.6    217  1  GTM3_MOUSE                P19639 mus musculu
   16     36    70.6    217  1  GTM3_RAT                  P08009 rattus norv
   17     36    70.6    217  1  GTM5_HUMAN                P46439 homo sapien
   18     36    70.6    217  1  GTMU_CAVPO                P16413 cavia porce
   19     36    70.6    217  1  GTMU_CRILO                Q00285 cricetulus
   20     36    70.6    217  1  GTMU_MESAU                P30116 mesocricetu
   21     36    70.6    217  1  GTMU_RABIT                P46409 oryctolagus
   22     36    70.6    218  1  GTM4_HUMAN                Q03013 homo sapien
   23     36    70.6    218  1  GTM6_MOUSE                O35660 mus musculu
   24     36    70.6    219  1  GTM2_CHICK                P20136 gallus gall
   25     36    70.6    224  1  GTM3_HUMAN                P21266 homo sapien
   26     36    70.6    224  1  GTM5_MOUSE                P48774 mus musculu
   27     36    70.6    226  1  BSH_DROME                 Q04787 drosophila
   28     36    70.6    358  1  PIAP_PIG                  O62640 sus scrofa
   29     36    70.6    370  1  TPS1_HUMAN                O60507 homo sapien
   30     36    70.6    370  1  TPS1_MOUSE                O70281 mus musculu
   31     36    70.6    381  1  RSA1_YEAST                Q08932 saccharomyc
   32     36    70.6    382  1  KSE1_ECOLI                P42501 escherichia
   33     36    70.6    382  1  KSE5_ECOLI                P42214 escherichia
```

```
   34     36    70.6    487  1  C8B1_ESCCA                O64899 eschscholzi
   35     36    70.6    488  1  C8B2_ESCCA                O64900 eschscholzi
   36     36    70.6    509  1  C4A1_DROME                Q9vms9 drosophila
   37     36    70.6    618  1  BIR3_HUMAN                Q13490 homo sapien
   38     35    68.6    273  1  TRY6_ANOGA                P35040 anopheles g
   39     35    68.6    380  1  TPSA_CAEEL                O77081 caenorhabdi
   40     35    68.6    500  1  C912_ARATH                O65790 arabidopsis
   41     34    66.7    180  1  YH95_AQUAE                O67664 aquifex aeo
   42     34    66.7    260  1  YH20_PASMU                Q9cka5 pasteurella
   43     34    66.7    855  1  ST14_MOUSE                P56677 mus musculu
   44     34    66.7   1097  1  CCT_DROME                 O96433 drosophila
   45     33    64.7    110  1  YLXM_STRMU                P96468 streptococc

                              ALIGNMENTS

RESULT 1
WT1_RAT
ID   WT1_RAT        STANDARD;      PRT;   448 AA.
AC   P49952;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Wilms' tumor protein homolog.
GN   WT1 OR WT-1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX   MEDLINE=93046155; PubMed=1330293;
RA   Sharma P.M., Yang X., Bowman M., Roberts V., Sukumar S.;
RT   "Molecular cloning of rat Wilms' tumor complementary DNA and a study
RT   of messenger RNA expression in the urogenital system and the brain.";
RL   Cancer Res. 52:6407-6412(1992).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; are
CC       produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: KIDNEY (BY SIMILARITY).
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING KIDNEY DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X69716; CAA49373.1; -.
DR   HSSP; P08046; 1AAY.
DR   TRANSFAC; T02352; -.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 1.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PRINTS; PR00048; ZINCFINGER.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 4.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene.
FT   DOMAIN       27     82       PRO-RICH.
FT   DOMAIN      322    437       ZINC FINGERS.
```

```
FT   ZN_FING     322    346       C2H2-TYPE.
FT   ZN_FING     352    376       C2H2-TYPE.
FT   ZN_FING     382    404       C2H2-TYPE.
FT   ZN_FING     413    437       C2H2-TYPE.
FT   VARSPLIC    249    265       MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT   VARSPLIC    407    409       MISSING (IN ISOFORM 2 AND ISOFORM 4).
SQ   SEQUENCE   448 AA;  49193 MW;  329AC9AC1FF73F76 CRC64;

  Query Match            100.0%;  Score 51;  DB 1;  Length 448;
  Best Local Similarity  100.0%;  Pred. No. 0.024;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     125 RMFPNAPYL 133


RESULT 2
WT1_HUMAN
ID   WT1_HUMAN      STANDARD;      PRT;   449 AA.
AC   P19544; Q16575;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Wilms' tumor protein (WT33).
GN   WT1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal kidney;
RX   MEDLINE=90158822; PubMed=2154702;
RA   Gessler M., Poustka A., Cavenee W., Neve R.L., Orkin S.H.,
RA   Bruns G.A.P.;
RT   "Homozygous deletion in Wilms tumours of a zinc-finger gene
RT   identified by chromosome jumping.";
RL   Nature 343:774-778(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=92052142; PubMed=1658787;
RA   Haber D.A., Sohn R.L., Buckler A.J., Pelletier J., Call K.M.,
RA   Housman D.E.;
RT   "Alternative splicing and genomic structure of the Wilms tumor gene
RT   WT1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:9618-9622(1991).
RN   [3]
RP   SEQUENCE OF 85-449 FROM N.A.
RX   MEDLINE=90150277; PubMed=2154335;
RA   Call K.M., Glaser T., Ito C.Y., Buckler A.J., Pelletier J.,
RA   Haber D.A., Rose E.A., Kral A., Yeger H., Lewis W.H., Jones C.,
RA   Housman D.E.;
RT   "Isolation and characterization of a zinc finger polypeptide gene at
RT   the human chromosome 11 Wilms' tumor locus.";
RL   Cell 60:509-520(1990).
RN   [4]
RP   IDENTIFICATION OF START CODON AND ALTERNATIVE SPLICE SITES.
RX   MEDLINE=91141522; PubMed=1671709;
RA   Buckler A.J., Pelletier J., Haber D.A., Glaser T., Housman D.E.;
RT   "Isolation, characterization, and expression of the murine Wilms'
RT   tumor gene (WT1) during kidney development.";
RL   Mol. Cell. Biol. 11:1707-1712(1991).
RN   [5]
RP   VARIANT WT CYS-366.
RX   MEDLINE=92279213; PubMed=1317572;
RA   Little M.H., Prosser J., Condie A., Smith P.J., van Heyningen V.,
RA   Hastie N.D.;
RT   "Zinc-finger point mutations within the WT1 gene in Wilms tumor
RT   patients.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:4791-4795(1992).
RN   [6]
```

```
RP   VARIANTS DDS.
RX   MEDLINE=92005721; PubMed=1655284;
RA   Pelletier J., Bruening W., Kashtan C.E., Mauer S.M., Manivel J.C.,
RA   Striegel J.E., Houghton D.C., Junien C., Habib R., Fouser L.,
RA   Fine R.N., Silverman B.L., Haber D.A., Housman D.;
RT   "Germline mutations in the Wilms' tumor suppressor gene are
RT   associated with abnormal urogenital development in Denys-Drash
RT   syndrome.";
RL   Cell 67:437-447(1991).
RN   [7]
RP   VARIANTS DDS.
RX   MEDLINE=93265053; PubMed=1338906;
RA   Baird P.N., Santos A., Groves N., Jadresic L., Cowell J.K.;
RT   "Constitutional mutations in the WT1 gene in patients with
RT   Denys-Drash syndrome.";
RL   Hum. Mol. Genet. 1:301-305(1992).
RN   [8]
RP   SEQUENCE OF 385-405 FROM N.A., AND VARIANT DDS TRP-394.
RX   MEDLINE=93250986; PubMed=1302008;
RA   Bruening W., Bardeesy N., Silverman B.L., Cohn R.A., Machin G.A.,
RA   Aronson A.J., Housman D., Pelletier J.;
RT   "Germline intronic and exonic mutations in the Wilms' tumour gene
RT   (WT1) affecting urogenital development.";
RL   Nat. Genet. 1:144-148(1992).
RN   [9]
RP   VARIANTS DDS.
RX   MEDLINE=93271983; PubMed=8388765;
RA   Little M.H., Williamson K.A., Mannens M., Kelsey A., Gosden C.,
RA   Hastie N.D., van Heyningen V.;
RT   "Evidence that WT1 mutations in Denys-Drash syndrome patients may act
RT   in a dominant-negative fashion.";
RL   Hum. Mol. Genet. 2:259-264(1993).
RN   [10]
RP   VARIANT MESOTHELIOMA GLY-273.
RX   MEDLINE=94004972; PubMed=8401592;
RA   Park S., Schalling M., Bernard A., Maheswaran S., Shipley G.C.,
RA   Roberts D., Fletcher J., Shipman R., Rheinwald J., Demetri G.,
RA   Griffin J., Minden M., Housman D.E., Haber D.A.;
RT   "The Wilms tumour gene WT1 is expressed in murine mesoderm-derived
RT   tissues and mutated in a human mesothelioma.";
RL   Nat. Genet. 4:415-420(1993).
RN   [11]
RP   VARIANTS WT SER-181 AND ALA-253.
RX   MEDLINE=97268681; PubMed=9108089;
RA   Schumacher V., Schneider S., Figge A., Wildhardt G., Harms D.,
RA   Schmidt D., Weirich A., Ludwig R., Royer-Pokora B.;
RT   "Correlation of germ-line mutations and two-hit inactivation of the
RT   WT1 gene with Wilms tumors of stromal-predominant histology.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:3972-3977(1997).
RN   [12]
RP   VARIANTS DMS TYR-377; LEU-383 AND ASN-396.
RX   MEDLINE=98198341; PubMed=9529364;
RA   Jeanpierre C., Denamur E., Henry I., Cabanis M.-O., Luce S.,
RA   Cecille A., Elion J., Peuchmaur M., Loirat C., Niaudet P.,
RA   Gubler M.-C., Junien C.;
RT   "Identification of constitutional WT1 mutations, in patients with
RT   isolated diffuse mesangial sclerosis, and analysis of
RT   genotype/phenotype correlations by use of a computerized mutation
RT   database.";
RL   Am. J. Hum. Genet. 62:824-833(1998).
RN   [13]
RP   REVIEW.
RX   MEDLINE=92207913; PubMed=1313285;
RA   Haber D.A., Buckler A.J.;
RT   "WT1: a novel tumor suppressor gene inactivated in Wilms' tumor.";
RL   New Biol. 4:97-106(1992).
RN   [14]
RP   REVIEW.
RX   MEDLINE=93345769; PubMed=8393820;
RA   Rauscher F.J. III;
RT   "The WT1 Wilms tumor gene product: a developmentally-regulated
RT   transcription factor in the kidney that functions as a tumor
RT   suppressor.";
```

```
RL   FASEB J. 7:896-903(1993).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; are
CC       produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN THE KIDNEY AND A SUBSET OF
CC       HEMATOPOIETIC CELLS.
CC   -!- DISEASE: Defects in WT1 are a cause of Wilm's tumor (WT), an
CC       embryonal malignancy of the kidney that affects approximately 1 in
CC       10'000 infants and young children. It occurs both in sporadic and
CC       hereditary forms.
CC   -!- DISEASE: Defects in WT1 are the cause of Denys-Drash syndrome
CC       (DDS), a disease characterized by a typical nephropathy and
CC       genital abnormalities.
CC   -!- DISEASE: Defects in WT1 are the cause of diffuse mesangial
CC       sclerosis (DMS), a form of Denys-Drash syndrome.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   -!- DATABASE: NAME=WT1; NOTE=WT1 mutation database;
CC       WWW="http://www.umd.necker.fr:2003/".
CC   -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC       WWW="http://www.infobiogen.fr/services/chromcancer/Genes/WT1ID78.html".
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X51630; CAA35956.1; ALT_INIT.
DR   EMBL; M80232; AAA61299.1; -.
DR   EMBL; M80217; AAA61299.1; JOINED.
DR   EMBL; M80218; AAA61299.1; JOINED.
DR   EMBL; M80219; AAA61299.1; JOINED.
DR   EMBL; M80220; AAA61299.1; JOINED.
DR   EMBL; M80221; AAA61299.1; JOINED.
DR   EMBL; M80228; AAA61299.1; JOINED.
DR   EMBL; M80229; AAA61299.1; JOINED.
DR   EMBL; M80231; AAA61299.1; JOINED.
DR   EMBL; M30393; AAA36810.1; -.
DR   EMBL; S61515; AAB20110.1; -.
DR   EMBL; S60755; AAC60605.1; -.
DR   PIR; A34673; A34673.
DR   PIR; S08273; S08273.
DR   HSSP; P08046; 1AAY.
DR   TRANSFAC; T00899; -.
DR   TRANSFAC; T00900; -.
DR   TRANSFAC; T01839; -.
DR   TRANSFAC; T01840; -.
DR   TRANSFAC; T01841; -.
DR   TRANSFAC; T01842; -.
DR   Genew; HGNC:12796; WT1.
DR   MIM; 194070; -.
DR   MIM; 194080; -.
DR   MIM; 256370; -.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 1.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PRINTS; PR00048; ZINCFINGER.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 4.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene;
KW   Disease mutation.
FT   DOMAIN       27     83       PRO-RICH.
FT   DOMAIN      323    438       ZINC FINGERS.
```

```
  Query Match             100.0%;  Score 51;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.025;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 RMFPNAPYL 134


RESULT 3
WT1_MOUSE
ID   WT1_MOUSE      STANDARD;      PRT;   449 AA.
AC   P22561;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Wilms' tumor protein homolog.
GN   WT1 OR WT-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91141522; PubMed=1671709;
RA   Buckler A.J., Pelletier J., Haber D.A., Glaser T., Housman D.E.;
RT   "Isolation, characterization, and expression of the murine Wilms'
RT   tumor gene (WT1) during kidney development.";
RL   Mol. Cell. Biol. 11:1707-1712(1991).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; are
CC       produced by alternative splicing.
CC   -!- TISSUE SPECIFICITY: KIDNEY.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING KIDNEY DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M55512; AAA40573.1; -.
DR   PIR; A39692; A39692.
DR   HSSP; P08046; 1AAY.
DR   TRANSFAC; T02351; -.
DR   MGD; MGI:98968; Wt1.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 1.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PRINTS; PR00048; ZINCFINGER.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 4.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene.
FT   DOMAIN       28     83       PRO-RICH.
FT   DOMAIN      323    438       ZINC FINGERS.
FT   ZN_FING     323    347       C2H2-TYPE.
FT   ZN_FING     353    377       C2H2-TYPE.
FT   ZN_FING     383    405       C2H2-TYPE.
FT   ZN_FING     414    438       C2H2-TYPE.
FT   VARSPLIC    250    266       MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT   VARSPLIC    408    410       MISSING (IN ISOFORM 2 AND ISOFORM 4).
```

```
SQ   SEQUENCE   449 AA;  49246 MW;  962381E9C8D7A380 CRC64;

  Query Match             100.0%;  Score 51;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.025;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 RMFPNAPYL 134


RESULT 4
WT1_PIG
ID   WT1_PIG        STANDARD;      PRT;   449 AA.
AC   O62651;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Wilms' tumor protein homolog.
GN   WT1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LWD; TISSUE=Kidney;
RX   MEDLINE=98267201; PubMed=9602131;
RA   Tsurutani N., Oda H., Nakatsuru Y., Imai Y., Zhang S., Ueno Y.,
RA   Ishikawa T.;
RT   "cDNA cloning and developmental expression of the porcine homologue
RT   of WT1.";
RL   Gene 211:215-220(1998).
CC   -!- FUNCTION: POTENTIAL ROLE IN TRANSCRIPTIONAL REGULATION. RECOGNIZES
CC       AND BINDS TO THE DNA SEQUENCE 5'-CGCCCCCGC-3'.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- ALTERNATIVE PRODUCTS: 4 isoforms; 1 (shown here), 2, 3 and 4; may
CC       be produced by alternative splicing.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED DURING KIDNEY DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE EGR FAMILY OF C2H2-TYPE ZINC-FINGER
CC       PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB010969; BAA28147.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   Pfam; PF02165; WT1; 1.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PRINTS; PR00048; ZINCFINGER.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 4.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein;
KW   Transcription regulation; Alternative splicing; Anti-oncogene.
FT   DOMAIN       28     83       PRO-RICH.
FT   DOMAIN      323    438       ZINC FINGERS.
FT   ZN_FING     323    347       C2H2-TYPE.
FT   ZN_FING     353    377       C2H2-TYPE.
FT   ZN_FING     383    405       C2H2-TYPE.
FT   ZN_FING     414    438       C2H2-TYPE.
FT   VARSPLIC    249    265       MISSING (IN ISOFORM 2 AND ISOFORM 3) (BY
FT                                SIMILARITY).
FT   VARSPLIC    407    409       MISSING (IN ISOFORM 2 AND ISOFORM 4) (BY
```

```
FT                                SIMILARITY).
SQ   SEQUENCE   449 AA;  49166 MW;  9C3E557B96F5A7B3 CRC64;

  Query Match             100.0%;  Score 51;  DB 1;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.025;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           |||||||||
Db     126 RMFPNAPYL 134


RESULT 5
YFIO_HAEIN
ID   YFIO_HAEIN     STANDARD;      PRT;   262 AA.
AC   P44553;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Putative lipoprotein HI0177 precursor.
GN   HI0177.
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC   Haemophilus.
OX   NCBI_TaxID=727;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rd / KW20 / ATCC 51907;
RX   MEDLINE=95350630; PubMed=7542800;
RA   Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA   Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA   McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA   Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA   Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA   Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA   Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA   Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA   Venter J.C.;
RT   "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT   Rd.";
RL   Science 269:496-512(1995).
RN   [2]
RP   IDENTIFICATION BY MASS SPECTROMETRY.
RX   MEDLINE=20137488; PubMed=10675023;
RA   Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
RA   Gray C., Fountoulakis M.;
RT   "Two-dimensional map of the proteome of Haemophilus influenzae.";
RL   Electrophoresis 21:411-429(2000).
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
CC       ANCHOR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE UPF0169 (COML) FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U32703; AAC21847.1; -.
DR   TIGR; HI0177; -.
DR   InterPro; IPR005156; UPF0169.
DR   Pfam; PF03696; UPF0169; 1.
KW   Membrane; Lipoprotein; Signal; Complete proteome.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    262       PUTATIVE LIPOPROTEIN HI0177.
FT   LIPID        19     19       N-ACYL DIGLYCERIDE (POTENTIAL).
SQ   SEQUENCE   262 AA;  29347 MW;  1DD9FFFE568D2B7B CRC64;

  Query Match              78.4%;  Score 40;  DB 1;  Length 262;
  Best Local Similarity    75.0%;  Pred. No. 1.7;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 RMFPNAPY 8
              |:|||:||
Db        151 RVFPNSPY 158


RESULT 6
IFT2_CRILO
ID   IFT2_CRILO     STANDARD;      PRT;   468 AA.
AC   Q60462;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Interferon-induced protein with tetratricopeptide repeats 2 (IFIT-2)
DE   (Interferon-induced 54 kDa protein) (IFI-54K) (CL-54 K).
GN   IFIT2 OR IFI54.
OS   Cricetulus longicaudatus (Long-tailed hamster) (Chinese hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10030;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94170787; PubMed=8125096;
RA   Bluyssen J.A.R., Vlietstra R.J., van der Made A., Trapman J.;
RT   "The interferon-stimulated gene 54 K promoter contains two adjacent
RT   functional interferon-stimulated response elements of different
RT   strength, which act synergistically for maximal interferon-alpha
RT   inducibility.";
RL   Eur. J. Biochem. 220:395-402(1994).
CC   -!- INDUCTION: BY INTERFERONS.
CC   -!- SIMILARITY: BELONGS TO THE IFIT FAMILY.
CC   -!- SIMILARITY: CONTAINS 6 TPR REPEATS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X77259; CAA54477.1; -.
DR   InterPro; IPR001440; TPR.
DR   Pfam; PF00515; TPR; 4.
KW   Repeat; TPR repeat; Interferon induction.
FT   REPEAT       94    127       TPR 1.
FT   REPEAT      138    171       TPR 2.
FT   REPEAT      175    208       TPR 3.
FT   REPEAT      242    275       TPR 4.
FT   REPEAT      327    360       TPR 5.
FT   REPEAT      364    398       TPR 6.
SQ   SEQUENCE   468 AA;  55045 MW;  9D92B878F81B1E27 CRC64;

  Query Match             72.5%;  Score 37;  DB 1;  Length 468;
  Best Local Similarity   66.7%;  Pred. No. 12;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              || ||:||:
Db        203 RMSPNSPYV 211


RESULT 7
UN33_CAEEL
ID   UN33_CAEEL     STANDARD;      PRT;   854 AA.
AC   Q01630;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Uncoordinated protein 33 (Protein unc-33).
GN   UNC-33.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RX   MEDLINE=93106371; PubMed=1468626;
RA   Li W., Herman R.K., Shaw J.E.;
RT   "Analysis of the Caenorhabditis elegans axonal guidance and outgrowth
RT   gene unc-33.";
RL   Genetics 132:675-689(1992).
CC   -!- FUNCTION: INVOLVED IN AXONAL GUIDANCE AND OUTGROWTH.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2/72.1 kDa
CC       and 3/55.5 kDa; may be produced by alternative splicing.
CC   -!- DEVELOPMENTAL STAGE: UNC-33 PROTEINS ARE DISTRIBUTED EXCLUSIVELY
CC       WITHIN NEURONAL PROCESSES AFTER EARLY EMBRYOGENESIS.
CC   -!- SIMILARITY: BELONGS TO THE DEHYDROPYRIMIDINASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z14148; CAA78520.1; -.
DR   EMBL; Z14148; CAA78521.1; -.
DR   EMBL; Z14148; CAA78522.1; -.
DR   EMBL; Z14146; CAA78516.1; -.
DR   EMBL; Z14146; CAA78517.1; -.
DR   EMBL; Z14146; CAA78518.1; -.
DR   PIR; S24643; S24643.
DR   PIR; S24644; S24644.
DR   PIR; S33558; S33558.
DR   InterPro; IPR002195; Dihydrooratase.
DR   Pfam; PF00744; Dihydroorotase; 1.
KW   Alternative splicing.
FT   VARSPLIC      1    175       MISSING (IN ISOFORM 2).
FT   VARSPLIC      1    331       MISSING (IN ISOFORM 3).
SQ   SEQUENCE   854 AA;  90819 MW;  A8073DDE251D2D77 CRC64;

  Query Match             72.5%;  Score 37;  DB 1;  Length 854;
  Best Local Similarity   66.7%;  Pred. No. 22;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |: ||:|||
Db        773 RLSPNSPYL 781


RESULT 8
BIR1_HUMAN
ID   BIR1_HUMAN     STANDARD;      PRT;  1403 AA.
AC   Q13075; Q13730; Q99796; O75857;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Baculoviral IAP repeat-containing protein 1 (Neuronal apoptosis
DE   inhibitory protein).
GN   BIRC1 OR NAIP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal brain;
RX   MEDLINE=95112344; PubMed=7813013;
RA   Roy N., Mahadevan M.S., McLean M., Shutler G., Yaraghi Z.,
RA   Farahini R., Baird S., Besner-Johnston A., Lefebvre C., Kang X.,
RA   Salih M., Aubry H., Tamai K., Guan X., Ioannou P., Crawford T.O.,
```

```
RA   de Jong P.J., Surh L., Ikeda J., Korneluk R.G., Mackenzie A.;
RT   "The gene for neuronal apoptosis inhibitory protein is partially
RT   deleted in individuals with spinal muscular atrophy.";
RL   Cell 80:167-178(1995).
RN   [2]
RP   SEQUENCE FROM N.A., AND REVISIONS.
RC   TISSUE=Brain;
RX   MEDLINE=98163755; PubMed=9503025;
RA   Chen Q., Baird S.D., Mahadevan M., Besner-Johnston A., Farahani R.,
RA   Xuan J.-Y., Kang X., Lefebvre C., Ikeda J.-E., Korneluk R.G.,
RA   MacKenzie A.E.;
RT   "Sequence of a 131-kb region of 5q13.1 containing the spinal muscular
RT   atrophy candidate genes SMN and NAIP.";
RL   Genomics 48:121-127(1998).
RN   [3]
RP   SEQUENCE OF 386-623 FROM N.A.
RA   der Steege G., Draaijers T.G., Grootscholten P.M., Osinga J.,
RA   Anzevino R., Velona I., Brahe C., Scheffer H., van Ommen G.J.B.,
RA   Buys C.H.C.M.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 222-1403 FROM N.A.
RA   Jones K., Graves T., McPherson J.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   FUNCTION.
RC   TISSUE=Liver;
RX   MEDLINE=96149249; PubMed=8552191;
RA   Liston P., Roy N., Tamai K., Lefebvre C., Baird S., Cherton-Horvat G.,
RA   Farahani R., McLean M., Ikeda J., Mackenzie A., Korneluk R.G.;
RT   "Suppression of apoptosis in mammalian cells by NAIP and a related
RT   family of IAP genes.";
RL   Nature 379:349-353(1996).
CC   -!- FUNCTION: PREVENTS MOTOR-NEURON APOPTOSIS INDUCED BY A VARIETY OF
CC       SIGNALS.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN MOTOR NEURONS, BUT NOT IN SENSORY
CC       NEURONS. FOUND IN LIVER AND PLACENTA, AND IN A LESSER EXTENT IN
CC       SPINAL CORD.
CC   -!- DISEASE: MUTATED OR DELETED FORMS OF NAIP HAVE BEEN FOUND IN
CC       INDIVIDUALS WITH SPINAL MUSCULAR ATROPHY TYPE I (SMA TYPE 1). SMAS
CC       ARE FATAL AUTOSOMAL RECESSIVE DISORDERS SUBCLASSIFIED AS TYPE 1
CC       (WERDNIG-HOFFMANN DISEASE), TYPE II (INTERMEDIATE FORM), AND TYPE
CC       III (WOHLFART-KUGELBERG-WELANDER DISEASE) BASED UPON THE AGE OF
CC       ONSET AND CLINICAL SEVERITY. THESE NEURODEGENERATIVE DISORDERS ARE
CC       CHARACTERIZED BY DEGENERATION OF LOWER MOTOR NEURONS, LEADING TO
CC       PROGRESSIVE PARALYSIS MUSCULAR ATROPHY. CONCERNS 1 IN 6000
CC       NEWBORNS.
CC   -!- SIMILARITY: CONTAINS 3 BIR REPEATS.
CC   -!- SIMILARITY: CONTAINS 1 NACHT DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U19251; AAC52045.1; -.
DR   EMBL; U80017; AAC52047.1; -.
DR   EMBL; U21913; AAA64504.1; -.
DR   EMBL; AC005031; AAC62261.1; -.
DR   HSSP; Q13490; 1QBH.
DR   Genew; HGNC:7634; BIRC1.
DR   MIM; 600355; -.
DR   InterPro; IPR001370; BIR.
DR   Pfam; PF00653; BIR; 3.
DR   SMART; SM00238; BIR; 3.
DR   PROSITE; PS01282; BIR_REPEAT_1; 3.
DR   PROSITE; PS50143; BIR_REPEAT_2; 3.
DR   PROSITE; PS50837; NACHT; 1.
KW   Apoptosis; Repeat.
FT   REPEAT       60    127       BIR 1.
FT   REPEAT      159    227       BIR 2.
FT   REPEAT      278    345       BIR 3.
FT   DOMAIN      464    758       NACHT.
FT   CONFLICT    222    223       PK -> YR (IN REF. 4).
FT   CONFLICT    386    387       VP -> ST (IN REF. 3).
FT   CONFLICT    535    535       M -> V (IN REF. 3).
FT   CONFLICT    553    553       Y -> H (IN REF. 3).
FT   CONFLICT   1228   1231       MISSING (IN REF. 4).
SQ   SEQUENCE   1403 AA;  159613 MW;  566304C154DA5E64 CRC64;

  Query Match             72.5%;  Score 37;  DB 1;  Length 1403;
  Best Local Similarity   66.7%;  Pred. No. 38;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            | ||| |:|
Db      337 RCFPNCPFL 345


RESULT 9
GTM1_HUMAN
ID   GTM1_HUMAN     STANDARD;      PRT;   217 AA.
AC   P09488;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Glutathione S-transferase Mu 1 (EC 2.5.1.18) (GSTM1-1) (HB subunit 4)
DE   (GTH4) (GSTM1a-1a) (GSTM1b-1b) (GST class-mu 1).
GN   GSTM1 OR GST1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88335606; PubMed=3419925;
RA   Dejong J.L., Chang C.M., Whang Peng J., Knutsen T., Tu C.-P.D.;
RT   "The human liver glutathione S-transferase gene superfamily:
RT   expression and chromosome mapping of an Hb subunit cDNA.";
RL   Nucleic Acids Res. 16:8541-8554(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89017184; PubMed=3174634;
RA   Seidegaard J., Vorachek W.R., Pero R.W., Pearson W.R.;
RT   "Hereditary differences in the expression of the human glutathione
RT   transferase active on trans-stilbene oxide are due to a gene
RT   deletion.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:7293-7297(1988).
RN   [3]
RP   SEQUENCE OF 1-188 FROM N.A.
RX   MEDLINE=93228631; PubMed=8471052;
RA   Zhong S., Spurr N.K., Hayes J.D., Wolf C.R.;
RT   "Deduced amino acid sequence, gene structure and chromosomal location
RT   of a novel human class Mu glutathione S-transferase, GSTM4.";
RL   Biochem. J. 291:41-50(1993).
RN   [4]
RP   SEQUENCE OF 59-117 FROM N.A.
RX   MEDLINE=90301515; PubMed=2362832;
RA   Comstock K.E., Sanderson B.J.S., Claflin G., Henner W.D.;
RT   "GST1 gene deletion determined by polymerase chain reaction.";
RL   Nucleic Acids Res. 18:3670-3670(1990).
RN   [5]
RP   SEQUENCE OF 1-23.
RX   MEDLINE=86042634; PubMed=3864155;
RA   Mannervik B., Alin P., Guthenberg C., Jensson H., Tahir M.K.,
RA   Warholm M., Joernvall H.;
RT   "Identification of three classes of cytosolic glutathione transferase
RT   common to several mammalian species: correlation between structural
RT   data and enzymatic properties.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:7202-7206(1985).
RN   [6]
RP   SEQUENCE OF 1-23.
RX   MEDLINE=85154554; PubMed=3979555;
```

```
RA   Alin P., Mannervik B., Joernvall H.;
RT   "Structural evidence for three different types of glutathione
RT   transferase in human tissues.";
RL   FEBS Lett. 182:319-322(1985).
RN   [7]
RP   SEQUENCE OF 52-59, AND MASS SPECTROMETRY.
RC   TISSUE=Liver;
RX   MEDLINE=21088920; PubMed=11271497;
RA   Hubbard M.J., McHugh N.J.;
RT   "Human ERp29: Isolation, primary structural characterisation and
RT   two-dimensional gel mapping.";
RL   Electrophoresis 21:3785-3796(2000).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.68 ANGSTROMS), AND MUTAGENESIS OF HIS-107.
RX   MEDLINE=99130249; PubMed=9930979;
RA   Patskovsky Y.V., Patskovska L.N., Listowsky I.;
RT   "Functions of His107 in the catalytic mechanism of human glutathione
RT   S-transferase hGSTM1a-1a.";
RL   Biochemistry 38:1193-1202(1999).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: THIS IS A LIVER ISOZYME.
CC   -!- POLYMORPHISM: THERE ARE TWO ALLELES; GSTM1A AND GSTM1B WHICH
CC       DIFFER IN POSITION 172.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X08020; CAA30821.1; -.
DR   EMBL; J03817; AAA59203.1; -.
DR   EMBL; X68676; CAA48636.1; -.
DR   EMBL; X51451; CAA35817.1; -.
DR   PIR; S01719; S01719.
DR   PIR; B22457; B22457.
DR   PIR; E24735; E24735.
DR   PDB; 1GTU; 02-FEB-99.
DR   Genew; HGNC:4632; GSTM1.
DR   MIM; 138350; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   InterPro; IPR003081; GST_mu.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family; Polymorphism; 3D-structure.
FT   INIT_MET      0      0
FT   VARIANT     172    172       K -> N (IN ALLELE B).
FT                                /FTId=VAR_003617.
FT   MUTAGEN     107    107       H->S: CHANGE THE PROPERTIES OF THE ENZMYE
FT                                TOWARD SOME SUBSTRATES.
FT   CONFLICT     43     43       S -> T (IN REF. 3).
SQ   SEQUENCE   217 AA;  25580 MW;  A8BB0E713BA75EAB CRC64;

  Query Match            70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 7.7;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        3 FPNAPYL 9
            ||| |||
Db       56 FPNLPYL 62

RESULT 10
GTM1_MOUSE
```

```
ID   GTM1_MOUSE     STANDARD;      PRT;   217 AA.
AC   P10649;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-JUL-1989 (Rel. 11, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Glutathione S-transferase Mu 1 (EC 2.5.1.18) (GST class-mu 1)
DE   (Glutathione S-transferase GT8.7) (pmGT10) (GST 1-1).
GN   GSTM1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88330838; PubMed=3417659;
RA   Pearson W.R., Reinhart J., Sisk S.C., Anderson K.S., Adler P.N.;
RT   "Tissue-specific induction of murine glutathione transferase mRNAs by
RT   butylated hydroxyanisole.";
RL   J. Biol. Chem. 263:13324-13332(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90094327; PubMed=2689439;
RA   Townsend A.J., Goldsmith M.E., Pickett C.B., Cowan K.H.;
RT   "Isolation, characterization, and expression in Escherichia coli of
RT   two murine Mu class glutathione S-transferase cDNAs homologous to the
RT   rat subunits 3 (Yb1) and 4 (Yb2).";
RL   J. Biol. Chem. 264:21582-21590(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93290350; PubMed=8512323;
RA   Reinhart J., Pearson W.R.;
RT   "The structure of two murine class-mu glutathione transferase genes
RT   coordinately induced by butylated hydroxyanisole.";
RL   Arch. Biochem. Biophys. 303:383-393(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Strausberg R.;
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   PRELIMINARY SEQUENCE OF 1-40.
RX   MEDLINE=83109018; PubMed=6822548;
RA   Pearson W.R., Windle J.J., Morrow J.F., Benson A.M., Talalay P.;
RT   "Increased synthesis of glutathione S-transferases in response to
RT   anticarcinogenic antioxidants. Cloning and measurement of messenger
RT   RNA.";
RL   J. Biol. Chem. 258:2052-2062(1983).
RN   [6]
RP   PRELIMINARY SEQUENCE OF 1-24.
RX   MEDLINE=86042634; PubMed=3864155;
RA   Mannervik B., Alin P., Guthenberg C., Jensson H., Tahir M.K.,
RA   Warholm M., Joernvall H.;
RT   "Identification of three classes of cytosolic glutathione transferase
RT   common to several mammalian species: correlation between structural
RT   data and enzymatic properties.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:7202-7206(1985).
RN   [7]
RP   CHARACTERIZATION.
RC   STRAIN=CD-1; TISSUE=Liver;
RX   MEDLINE=96189427; PubMed=8605288;
RA   Mitchell A.E., Morin D., Lame M.W., Jones A.D.;
RT   "Purification, mass spectrometric characterization, and covalent
RT   modification of murine glutathione S-transferases.";
RL   Chem. Res. Toxicol. 8:1054-1062(1995).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MASS SPECTROMETRY: MW=25838.4; MW_ERR=2; METHOD=Electrospray.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
```

```
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; J03952; AAA37747.1; -.
DR   EMBL; J04632; AAA37705.1; -.
DR   EMBL; L13448; -; NOT_ANNOTATED_CDS.
DR   EMBL; BC003822; AAH03822.1; -.
DR   PIR; A20831; A20831.
DR   PIR; A28946; A28946.
DR   PIR; I24735; I24735.
DR   PIR; A34159; A34159.
DR   PIR; S33860; S33860.
DR   HSSP; P04905; 2GST.
DR   SWISS-2DPAGE; P10649; MOUSE.
DR   MGD; MGI:95860; Gstm1.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   InterPro; IPR003081; GST_mu.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family.
FT   INIT_MET      0      0
SQ   SEQUENCE   217 AA;  25839 MW;  A1EE3938F590B829 CRC64;

  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 7.7;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        3 FPNAPYL 9
            ||| |||
Db       56 FPNLPYL 62


RESULT 11
GTM1_RAT
ID   GTM1_RAT       STANDARD;      PRT;   217 AA.
AC   P04905;
DT   13-AUG-1987 (Rel. 05, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glutathione S-transferase Yb1 (EC 2.5.1.18) (Chain 3) (GST M1-1)
DE   (GST class-mu 1).
GN   GSTM1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (CLONE PGTR200).
RX   MEDLINE=86312882; PubMed=2875437;
RA   Lai H.-C.J., Grove G., Tu C.-P.D.;
RT   "Cloning and sequence analysis of a cDNA for a rat liver glutathione
RT   S-transferase Yb subunit.";
RL   Nucleic Acids Res. 14:6101-6114(1986).
RN   [2]
RP   SEQUENCE FROM N.A. (CLONE PGTA/C44).
RX   MEDLINE=86033768; PubMed=3840477;
RA   Ding G.J.-F., Lu A.Y.H., Pickett C.B.;
RT   "Rat liver glutathione S-transferases. Nucleotide sequence analysis
RT   of a Yb1 cDNA clone and prediction of the complete amino acid
RT   sequence of the Yb1 subunit.";
RL   J. Biol. Chem. 260:13268-13271(1985).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86224097; PubMed=3011803;
RA   Ding G.J.-F., Ding V.D.-H., Rodkey J.A., Bennett C.D., Lu A.Y.H.,
RA   Pickett C.B.;
RT   "Rat liver glutathione S-transferases. DNA sequence analysis of a Yb2
RT   cDNA clone and regulation of the Yb1 and Yb2 mRNAs by
```

```
RT   phenobarbital.";
RL   J. Biol. Chem. 261:7952-7957(1986).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87308179; PubMed=3040722;
RA   Chang C., Saltzman A.G., Sorensen N.S., Hiipakka R.A., Liao S.;
RT   "Identification of glutathione S-transferase Yb1 mRNA as the
RT   androgen-repressed mRNA by cDNA cloning and sequence analysis.";
RL   J. Biol. Chem. 262:11901-11903(1987).
RN   [5]
RP   SEQUENCE OF 1-23.
RC   STRAIN=Wistar; TISSUE=Olfactory epithelium;
RX   MEDLINE=93277499; PubMed=8503873;
RA   Ben-Arie N., Khen M., Lancet D.;
RT   "Glutathione S-transferases in rat olfactory epithelium:
RT   purification, molecular properties and odorant biotransformation.";
RL   Biochem. J. 292:379-384(1993).
RN   [6]
RP   MUTAGENESIS OF CYS-86.
RX   MEDLINE=91354218; PubMed=1883338;
RA   Hsieh J.-C., Huang S.-C., Chen W.-L., Lai Y.-C., Tam M.F.;
RT   "Cysteine-86 is not needed for the enzymic activity of glutathione S-
RT   transferase 3-3.";
RL   Biochem. J. 278:293-297(1991).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX   MEDLINE=93041702; PubMed=1420139;
RA   Ji X., Zhang P., Armstrong R.N., Gilliland G.L.;
RT   "The three-dimensional structure of a glutathione S-transferase from
RT   the mu gene class. Structural analysis of the binary complex of
RT   isoenzyme 3-3 and glutathione at 2.2-A resolution.";
RL   Biochemistry 31:10169-10184(1992).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RA   Fu J.-H., Rose J., Tam M.F., Wang B.-C.;
RT   "New crystal forms of a mu-class glutathione S-transferase from rat
RT   liver.";
RL   Acta Crystallogr. D 50:219-224(1994).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX   MEDLINE=94153886; PubMed=8110735;
RA   Ji X., Johnson W.W., Sesay M.A., Dickert L., Prasad S.M., Ammon H.L.,
RA   Armstrong R.N., Gilliland G.L.;
RT   "Structure and function of the xenobiotic substrate binding site of a
RT   glutathione S-transferase as revealed by X-ray crystallographic
RT   analysis of product complexes with the diastereomers of 9-(S-
RT   glutathionyl)-10-hydroxy-9,10-dihydrophenanthrene.";
RL   Biochemistry 33:1043-1052(1994).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- FUNCTION: THE OLFACTORY GST MAY BE CRUCIAL FOR THE ACUITY OF THE
CC       OLFACTORY PROCESS.
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- SUBUNIT: HOMODIMER OR HETERODIMER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: RAT LIVER GLUTATHIONE S-TRANSFERASES (GST) CONSIST
CC       OF BINARY COMBINATIONS OF THREE MAJOR CLASSES OF SUBUNITS
CC       DESIGNATED: YA (25 KDA), YB (27 KDA) AND YC (28 KDA).
CC   -!- MISCELLANEOUS: YB SUBCLASS SELECTIVELY BINDS STEROID HORMONES.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X04229; CAA27811.1; -.
DR   EMBL; M11719; AAA41287.1; -.
DR   EMBL; J02810; AAA41293.1; -.
DR   PIR; A24085; A24085.
```

```
DR   PIR; A25510; A25510.
DR   PIR; A29794; A29794.
DR   PIR; S17167; S17167.
DR   PDB; 1GSB; 31-OCT-93.
DR   PDB; 1GSC; 31-OCT-93.
DR   PDB; 2GST; 31-OCT-93.
DR   PDB; 3GST; 31-JAN-94.
DR   PDB; 4GST; 31-OCT-93.
DR   PDB; 5GST; 31-OCT-93.
DR   PDB; 6GST; 08-NOV-96.
DR   PDB; 6GSU; 08-NOV-96.
DR   PDB; 6GSV; 08-NOV-96.
DR   PDB; 6GSW; 08-NOV-96.
DR   PDB; 6GSX; 08-NOV-96.
DR   PDB; 6GSY; 08-NOV-96.
DR   PDB; 5FWG; 27-JAN-99.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   InterPro; IPR003081; GST_mu.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family; 3D-structure; Olfaction.
FT   INIT_MET      0      0
FT   MUTAGEN      86     86       C->S: NO CHANGE IN ACTIVITY.
FT   CONFLICT    168    168       I -> N (IN REF. 3).
FT   CONFLICT    198    199       KS -> NC (IN REF. 2).
FT   STRAND        2      7
FT   TURN         11     13
FT   HELIX        14     22
FT   TURN         23     24
FT   STRAND       27     32
FT   TURN         37     39
FT   HELIX        43     46
FT   TURN         47     50
FT   STRAND       61     64
FT   TURN         65     66
FT   STRAND       67     70
FT   HELIX        72     82
FT   TURN         83     84
FT   HELIX        90    114
FT   TURN        115    115
FT   TURN        117    118
FT   HELIX       119    128
FT   TURN        129    129
FT   HELIX       130    141
FT   TURN        142    143
FT   TURN        154    154
FT   HELIX       155    169
FT   TURN        171    176
FT   HELIX       178    188
FT   TURN        189    189
FT   HELIX       191    196
FT   TURN        197    198
FT   TURN        200    201
FT   TURN        210    211
SQ   SEQUENCE   217 AA;  25782 MW;  2ACE8D49DA785118 CRC64;

  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 7.7;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           ||| |||
Db      56 FPNLPYL 62


RESULT 12
GTM2_HUMAN
ID   GTM2_HUMAN     STANDARD;      PRT;   217 AA.
AC   P28161;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Glutathione S-transferase Mu 2 (EC 2.5.1.18) (GSTM2-2) (GST class-mu
DE   2).
GN   GSTM2 OR GST4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91239584; PubMed=2034681;
RA   Vorachek W.R., Pearson W.R., Rule G.S.;
RT   "Cloning, expression, and characterization of a class-mu glutathione
RT   transferase from human muscle, the product of the GST4 locus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:4443-4447(1991).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RX   MEDLINE=94238693; PubMed=8182750;
RA   Raghunathan S., Chandross R.J., Kretsinger R.H., Allison T.J.,
RA   Penington C.J., Rule G.S.;
RT   "Crystal structure of human class mu glutathione transferase GSTM2-2.
RT   Effects of lattice packing on conformational heterogeneity.";
RL   J. Mol. Biol. 238:815-832(1994).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- TISSUE SPECIFICITY: THIS IS A MUSCLE ISOZYME.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M63509; AAA60963.1; -.
DR   PIR; A39375; A39375.
DR   PDB; 1HNA; 31-JAN-94.
DR   PDB; 1HNB; 31-JAN-94.
DR   PDB; 1HNC; 31-JAN-94.
DR   PDB; 2GTU; 02-MAR-99.
DR   PDB; 3GTU; 29-JUL-99.
DR   Genew; HGNC:4634; GSTM2.
DR   MIM; 138380; -.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   InterPro; IPR003081; GST_mu.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family; 3D-structure.
FT   INIT_MET      0      0
FT   STRAND        2      5
FT   STRAND        7      7
FT   TURN         11     12
FT   HELIX        13     22
FT   TURN         23     24
FT   STRAND       27     29
FT   STRAND       32     32
FT   STRAND       35     35
FT   TURN         38     39
FT   STRAND       41     41
FT   HELIX        43     49
FT   TURN         50     52
FT   STRAND       61     64
FT   TURN         65     66
FT   STRAND       67     70
FT   HELIX        72     81
FT   TURN         82     84
```

```
FT   HELIX        90    113
FT   TURN        114    115
FT   TURN        117    118
FT   HELIX       119    127
FT   TURN        128    129
FT   HELIX       130    141
FT   TURN        142    143
FT   STRAND      146    146
FT   TURN        147    147
FT   STRAND      148    149
FT   TURN        150    150
FT   STRAND      151    151
FT   HELIX       154    169
FT   HELIX       171    176
FT   HELIX       178    189
FT   HELIX       191    197
FT   TURN        198    198
FT   TURN        214    215
SQ   SEQUENCE   217 AA;  25613 MW;  25603A909482CA39 CRC64;

  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 7.7;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        3 FPNAPYL 9
            ||| |||
Db       56 FPNLPYL 62


RESULT 13
GTM2_MOUSE
ID   GTM2_MOUSE     STANDARD;      PRT;   217 AA.
AC   P15626;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Glutathione S-transferase Mu 2 (EC 2.5.1.18) (GST class-mu 2)
DE   (Glutathione S-transferase pmGT2) (GST 5-5).
GN   GSTM2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90094327; PubMed=2689439;
RA   Townsend A.J., Goldsmith M.E., Pickett C.B., Cowna K.H.;
RT   "Isolation, characterization, and expression in Escherichia coli of
RT   two murine Mu class glutathione S-transferase cDNAs homologous to the
RT   rat subunits 3 (Yb1) and 4 (Yb2).";
RL   J. Biol. Chem. 264:21582-21590(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RX   MEDLINE=21297196; PubMed=11404019;
RA   Kumar A., Reddy E.P.;
RT   "Genomic organization and characterization of the promoter region of
RT   murine GSTM2 gene.";
RL   Gene 270:221-229(2001).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J04696; AAA37706.1; -.
DR   EMBL; AF319526; AAK28508.1; -.
DR   PIR; B34159; B34159.
DR   HSSP; P09488; 1GTU.
DR   SWISS-2DPAGE; P15626; MOUSE.
DR   MGD; MGI:95861; Gstm2.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   InterPro; IPR003081; GST_mu.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family.
FT   INIT_MET      0      0
SQ   SEQUENCE   217 AA;  25585 MW;  A3979D4435EF08AE CRC64;

  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 7.7;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        3 FPNAPYL 9
            ||| |||
Db       56 FPNLPYL 62


RESULT 14
GTM2_RAT
ID   GTM2_RAT       STANDARD;      PRT;   217 AA.
AC   P08010;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Glutathione S-transferase Yb2 (EC 2.5.1.18) (Chain 4) (GST class-mu
DE   2).
GN   GSTM2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88298790; PubMed=3403534;
RA   Lai H.-C.J., Qian B., Grove G., Tu C.-P.D.;
RT   "Gene expression of rat glutathione S-transferases. Evidence for gene
RT   conversion in the evolution of the Yb multigene family.";
RL   J. Biol. Chem. 263:11389-11395(1988).
RN   [2]
RP   SEQUENCE.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=86192461; PubMed=3699019;
RA   Alin P., Mannervik B., Joernvall H.;
RT   "Cytosolic rat liver glutathione transferase 4-4. Primary structure
RT   of the protein reveals extensive differences between homologous
RT   glutathione transferases of classes alpha and mu.";
RL   Eur. J. Biochem. 156:343-350(1986).
RN   [3]
RP   SEQUENCE OF 24-217 FROM N.A.
RX   MEDLINE=86224097; PubMed=3011803;
RA   Ding G.J.-F., Ding V.D.-H., Rodkey J.A., Bennett C.D., Lu A.Y.H.,
RA   Pickett C.B.;
RT   "Rat liver glutathione S-transferases. DNA sequence analysis of a Yb2
RT   cDNA clone and regulation of the Yb1 and Yb2 mRNAs by
RT   phenobarbital.";
RL   J. Biol. Chem. 261:7952-7957(1986).
RN   [4]
RP   SEQUENCE OF 32-217 FROM N.A.
RX   MEDLINE=87008619; PubMed=3020050;
RA   Lai H.-C.J., Tu C.-P.D.;
RT   "Rat glutathione S-transferases supergene family. Characterization of
RT   an anionic Yb subunit cDNA clone.";
RL   J. Biol. Chem. 261:13793-13799(1986).
RN   [5]
```

```
RP   SEQUENCE OF 1-21.
RC   STRAIN=Wistar; TISSUE=Olfactory epithelium;
RX   MEDLINE=93277499; PubMed=8503873;
RA   Ben-Arie N., Khen M., Lancet D.;
RT   "Glutathione S-transferases in rat olfactory epithelium:
RT   purification, molecular properties and odorant biotransformation.";
RL   Biochem. J. 292:379-384(1993).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- FUNCTION: THE OLFACTORY GST MAY BE CRUCIAL FOR THE ACUITY OF THE
CC       OLFACTORY PROCESS.
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- SUBUNIT: HOMODIMER OR HETERODIMER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: RAT LIVER GLUTATHIONE S-TRANSFERASES (GST) CONSIST
CC       OF BINARY COMBINATIONS OF THREE MAJOR CLASSES OF SUBUNITS
CC       DESIGNATED: YA (25 KDA), YB (27 KDA) AND YC (28 KDA).
CC   -!- MISCELLANEOUS: YB SUBCLASS SELECTIVELY BINDS STEROID HORMONES.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J02592; AAA41285.1; -.
DR   EMBL; M13590; AAA42351.1; -.
DR   EMBL; J03914; AAA41296.1; -.
DR   PIR; A25386; XURTG4.
DR   PIR; A26307; A26307.
DR   PIR; B26187; B26187.
DR   HSSP; P09488; 1GTU.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   InterPro; IPR003081; GST_mu.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family; Olfaction.
FT   INIT_MET      0      0
FT   CONFLICT    146    146       W -> S (IN REF. 2).
SQ   SEQUENCE   217 AA;  25571 MW;  F27B3D5831FF789F CRC64;

  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 7.7;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 FPNAPYL 9
              ||| |||
Db         56 FPNLPYL 62


RESULT 15
GTM3_MOUSE
ID   GTM3_MOUSE     STANDARD;      PRT;   217 AA.
AC   P19639;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Glutathione S-transferase Mu 3 (EC 2.5.1.18) (GST class-mu 3)
DE   (Glutathione S-transferase GT9.3).
GN   GSTM3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88330838; PubMed=3417659;
RA   Pearson W.R., Reinhart J., Sisk S.C., Anderson K.S., Adler P.N.;
RT   "Tissue-specific induction of murine glutathione transferase mRNAs by
RT   butylated hydroxyanisole.";
RL   J. Biol. Chem. 263:13324-13332(1988).
RN   [2]
RP   SEQUENCE OF 1-40 FROM N.A.
RX   MEDLINE=83109018; PubMed=6822548;
RA   Pearson W.R., Windle J.J., Morrow J.F., Benson A.M., Talalay P.;
RT   "Increased synthesis of glutathione S-transferases in response to
RT   anticarcinogenic antioxidants. Cloning and measurement of messenger
RT   RNA.";
RL   J. Biol. Chem. 258:2052-2062(1983).
RN   [3]
RP   SEQUENCE OF 1-25.
RX   MEDLINE=86042634; PubMed=3864155;
RA   Mannervik B., Alin P., Guthenberg C., Jensson H., Tahir M.K.,
RA   Warholm M., Joernvall H.;
RT   "Identification of three classes of cytosolic glutathione transferase
RT   common to several mammalian species: correlation between structural
RT   data and enzymatic properties.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:7202-7206(1985).
CC   -!- FUNCTION: CONJUGATION OF REDUCED GLUTATHIONE TO A WIDE NUMBER
CC       OF EXOGENOUS AND ENDOGENOUS HYDROPHOBIC ELECTROPHILES.
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: BELONGS TO THE GST SUPERFAMILY. MU FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J03953; AAA37748.1; ALT_INIT.
DR   PIR; B20831; B20831.
DR   PIR; B28946; B28946.
DR   PIR; E37520; E37520.
DR   HSSP; P04905; 2GST.
DR   MGD; MGI:106026; Gstm3.
DR   InterPro; IPR004046; GST_Cterm.
DR   InterPro; IPR004045; GST_Nterm.
DR   InterPro; IPR003081; GST_mu.
DR   Pfam; PF00043; GST_C; 1.
DR   Pfam; PF02798; GST_N; 1.
DR   PRINTS; PR01267; GSTRNSFRASEM.
KW   Transferase; Multigene family.
FT   INIT_MET      0      0
SQ   SEQUENCE   217 AA;  25570 MW;  C4D8950EFEA585D8 CRC64;

  Query Match             70.6%;  Score 36;  DB 1;  Length 217;
  Best Local Similarity   85.7%;  Pred. No. 7.7;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 FPNAPYL 9
              ||| |||
Db         56 FPNLPYL 62


Search completed: April 25, 2003, 06:46:36
Job time : 13 secs
```

```
                         GenCore version 5.1.4_p5_4578
                  Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:         April 25, 2003, 06:36:22 ; Search time 29 Seconds
                                           (without alignments)
                                           63.946 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters:      671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_21:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_virus:*
                13:  sp_vertebrate:*
                14:  sp_unclassified:*
                15:  sp_rvirus:*
                16:  sp_bacteriap:*
                17:  sp_archeap:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                   Description
-------------------------------------------------------------------------
    1      51   100.0    407  13  P79958               P79958 xenopus lae
    2      51   100.0    446  4  Q15881               Q15881 homo sapien
    3      47    92.2    390  13  O42223               O42223 trachemys s
    4      43    84.3    386  13  Q90XX8               Q90xx8 oncorhynchu
    5      43    84.3    388  13  Q90XX7               Q90xx7 oncorhynchu
    6      43    84.3    409  13  Q91657               Q91657 xenopus lae
    7      43    84.3    417  13  Q90XX6               Q90xx6 oncorhynchu
    8      41    80.4    213  16  Q929M3               Q929m3 listeria in
    9      41    80.4    213  16  Q8Y5B7               Q8y5b7 listeria mo
   10      40    78.4    250  13  Q98SJ4               Q98sj4 brachydanio
   11      40    78.4    289  16  Q8XPG7               Q8xpg7 ralstonia s
   12      40    78.4    392  13  Q91030               Q91030 gallus gall
   13      40    78.4    414  13  Q9I8A1               Q9i8a1 gallus gall
   14      40    78.4    417  13  Q9I8A0               Q9i8a0 gallus gall
   15      40    78.4    419  13  Q9PUT7               Q9put7 brachydanio
   16      39    76.5    203  16  O83076               O83076 treponema p
   17      39    76.5    392  13  Q9IBF0               Q9ibf0 anguilla ja
   18      39    76.5    426  13  Q9W611               Q9w611 cynops pyrr
   19      38    74.5    205  16  Q9I4D6               Q9i4d6 pseudomonas
   20      38    74.5    214  16  Q8YSH2               Q8ysh2 anabaena sp
   21      38    74.5    581  10  Q9FZ97               Q9fz97 arabidopsis
   22      38    74.5    581  10  Q8W4K7               Q8w4k7 arabidopsis
   23      38    74.5    662  5  Q9N6S5               Q9n6s5 leishmania
   24      38    74.5    956  5  Q961M6               Q961m6 drosophila
   25      38    74.5   2175  5  Q9W198               Q9w198 drosophila
   26      37    72.5    444  10  Q9LS41               Q9ls41 arabidopsis
   27      37    72.5    529  5  Q966C1               Q966c1 caenorhabdi
   28      37    72.5    541  3  Q8TG25               Q8tg25 aspergillus
   29      37    72.5   1160  4  Q8TDZ4               Q8tdz4 homo sapien
   30      36    70.6     32  11  Q9QVM0               Q9qvm0 mus sp. glu
   31      36    70.6     37  11  Q9QVM3               Q9qvm3 mus sp. glu
   32      36    70.6     82  4  Q9UE37               Q9ue37 homo sapien
   33      36    70.6    118  4  Q96BP2               Q96bp2 homo sapien
   34      36    70.6    181  4  Q8TC98               Q8tc98 homo sapien
   35      36    70.6    188  6  Q9MZB4               Q9mzb4 capra hircu
   36      36    70.6    195  4  Q05465               Q05465 homo sapien
   37      36    70.6    218  6  Q9N0V4               Q9n0v4 bos taurus
   38      36    70.6    218  6  Q9TSM5               Q9tsm5 macaca fasc
   39      36    70.6    218  6  Q9TSM4               Q9tsm4 macaca fasc
   40      36    70.6    218  11  Q9DD25               Q9dd25 mus musculu
   41      36    70.6    218  11  Q9DCE8               Q9dce8 mus musculu
   42      36    70.6    218  11  Q91Y83               Q91y83 cavia porce
   43      36    70.6    218  11  Q8R5I6               Q8r5i6 mus musculu
   44      36    70.6    219  13  Q90WM9               Q90wm9 xenopus lae
   45      36    70.6    220  4  Q8WWE1               Q8wwe1 homo sapien


                                ALIGNMENTS


RESULT 1
P79958
ID   P79958      PRELIMINARY;      PRT;   407 AA.
AC   P79958;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   WT1 protein.
GN   WT1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RX   MEDLINE=97074667; PubMed=8917094;
RA   Semba K., Saito-Ueno R., Takayama G., Kondo M.;
RT   "cDNA clonig and its pronephros-specific expression of the Wilms'
RT   tumor suppressor gene, WT1, from Xenopus laevis.";
RL   Gene 175:167-172(1996).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; D82051; BAA11522.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF02165; WT1; 1.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00048; ZINCFINGER.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 4.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   407 AA;  45983 MW;  E2554C658005870C CRC64;

  Query Match             100.0%;  Score 51;  DB 13;  Length 407;
  Best Local Similarity   100.0%;  Pred. No. 0.085;
```

```
  Matches     9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db      107 RMFPNAPYL 115


RESULT 2
Q15881
ID   Q15881      PRELIMINARY;      PRT;   446 AA.
AC   Q15881;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Wilms tumor gene 1, exon 1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92241883; PubMed=1572653;
RA   Gessler M., Konig A., Bruns G.A.;
RT   "The genomic organization and expression of the WT1 gene.";
RL   Genomics 12:807-813(1992).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; X61631; CAA43819.1; -.
DR   EMBL; X61632; CAA43819.1; JOINED.
DR   EMBL; X61633; CAA43819.1; JOINED.
DR   EMBL; X61634; CAA43819.1; JOINED.
DR   EMBL; X61635; CAA43819.1; JOINED.
DR   EMBL; X61636; CAA43819.1; JOINED.
DR   EMBL; X61637; CAA43819.1; JOINED.
DR   EMBL; X61638; CAA43819.1; JOINED.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF02165; WT1; 1.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   PRINTS; PR00048; ZINCFINGER.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 4.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   446 AA;  48842 MW;  8CE7FC047F41CF11 CRC64;

  Query Match            100.0%;  Score 51;  DB 4;  Length 446;
  Best Local Similarity  100.0%;  Pred. No. 0.093;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db      126 RMFPNAPYL 134


RESULT 3
O42223
ID   O42223      PRELIMINARY;      PRT;   390 AA.
AC   O42223;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Wilms tumor 1 protein.
GN   TWT1.
OS   Trachemys scripta (Red-eared slider turtle) (Pseudemys scripta).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Testudines; Cryptodira; Testudinoidea; Emydidae; Trachemys.
OX   NCBI_TaxID=34903;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Spotila L.D., Hall S.E.;
RT   "Expression of a new RNA-splice isoform of WT1 in developing
RT   kidney/gonadal complexes of the turtle, Trachemys scripta.";
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; AF019779; AAB70832.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF02165; WT1; 1.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00048; ZINCFINGER.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 4.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   390 AA;  43620 MW;  1C9987435BE927C2 CRC64;

  Query Match             92.2%;  Score 47;  DB 13;  Length 390;
  Best Local Similarity   88.9%;  Pred. No. 0.48;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            ||||| |||
Db      109 RMFPNGPYL 117


RESULT 4
Q90XX8
ID   Q90XX8      PRELIMINARY;      PRT;   386 AA.
AC   Q90XX8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Wilms' tumor suppressor 1a.
GN   WT-T1A.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21337607; PubMed=11444705;
RA   Brunelli J.P., Robison B.D., Thorgaard G.H.;
RT   "Ancient and recent duplications of the rainbow trout Wilms' tumor
RT   gene.";
RL   Genome 44:455-462(2001).
DR   EMBL; AF334670; AAK52719.1; -.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF02165; WT1; 1.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Zinc-finger.
SQ   SEQUENCE   386 AA;  43286 MW;  05622EF7642FC9E5 CRC64;

  Query Match             84.3%;  Score 43;  DB 13;  Length 386;
  Best Local Similarity   88.9%;  Pred. No. 2.8;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            ||| |||||
Db      112 RMFSNAPYL 120


RESULT 5
Q90XX7
ID   Q90XX7      PRELIMINARY;      PRT;   388 AA.
AC   Q90XX7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
```

```
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Wilms' tumor suppressor 1b.
GN   WT-T1B.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21337607; PubMed=11444705;
RA   Brunelli J.P., Robison B.D., Thorgaard G.H.;
RT   "Ancient and recent duplications of the rainbow trout Wilms' tumor
RT   gene.";
RL   Genome 44:455-462(2001).
DR   EMBL; AF334671; AAK52720.1; -.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF02165; WT1; 1.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Zinc-finger.
SQ   SEQUENCE   388 AA;  43398 MW;  BC92266DE745DFCC CRC64;

  Query Match             84.3%;  Score 43;  DB 13;  Length 388;
  Best Local Similarity   88.9%;  Pred. No. 2.8;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           ||| |||||
Db     112 RMFSNAPYL 120


RESULT 6
Q91657
ID   Q91657      PRELIMINARY;      PRT;   409 AA.
AC   Q91657;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   WILMS' tumor suppressor (WT1).
GN   WT1.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Carroll T.J., Vize P.D.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 289-370 FROM N.A.
RC   TISSUE=MESONEPHROS;
RX   MEDLINE=96068905; PubMed=7478606;
RA   Kent J., Coriat A.M., Sharpe P.T., Hastie N., Van Heyningen V.;
RT   "The evolution of WT1 sequence and expression pattern in the
RT   vertebrates.";
RL   Oncogene 11:1781-1792(1995).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; U42011; AAB53152.1; -.
DR   EMBL; X85733; CAA59738.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF02165; WT1; 1.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00048; ZINCFINGER.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 4.
```

```
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc-finger.
SQ   SEQUENCE   409 AA;  46186 MW;  2217FC04612CDE10 CRC64;

  Query Match             84.3%;  Score 43;  DB 13;  Length 409;
  Best Local Similarity   88.9%;  Pred. No. 3;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           ||| |||||
Db     106 RMFSNAPYL 114


RESULT 7
Q90XX6
ID   Q90XX6      PRELIMINARY;      PRT;   417 AA.
AC   Q90XX6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Wilms' tumor suppressor 1b variant.
GN   WT-T1B.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21337607; PubMed=11444705;
RA   Brunelli J.P., Robison B.D., Thorgaard G.H.;
RT   "Ancient and recent duplications of the rainbow trout Wilms' tumor
RT   gene.";
RL   Genome 44:455-462(2001).
DR   EMBL; AF334672; AAK52721.1; -.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF02165; WT1; 1.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Zinc-finger.
SQ   SEQUENCE   417 AA;  46557 MW;  5217620CD471D030 CRC64;

  Query Match             84.3%;  Score 43;  DB 13;  Length 417;
  Best Local Similarity   88.9%;  Pred. No. 3;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           ||| |||||
Db     112 RMFSNAPYL 120


RESULT 8
Q929M3
ID   Q929M3      PRELIMINARY;      PRT;   213 AA.
AC   Q929M3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical protein lin2252.
GN   LIN2252.
OS   Listeria innocua.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Listeriaceae; Listeria.
OX   NCBI_TaxID=1642;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CLIP 11262 / SEROVAR 6A;
RX   PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
```

```
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL596171; CAC97480.1; -.
DR   ListiList; LIN02252; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   213 AA;  25341 MW;  BB7F90FD1D1F0E78 CRC64;

  Query Match             80.4%;  Score 41;  DB 16;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           |||||||
Db      67 FPNAPYL 73


RESULT 9
Q8Y5B7
ID   Q8Y5B7      PRELIMINARY;      PRT;   213 AA.
AC   Q8Y5B7;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical protein lmo2148.
GN   LMO2148.
OS   Listeria monocytogenes.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC   Listeriaceae; Listeria.
OX   NCBI_TaxID=1639;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EGD-E / SEROVAR 1/2A;
RX   MEDLINE=21537279; PubMed=11679669;
RA   Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA   Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
RA   Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA   Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA   Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA   Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA   Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
RA   Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA   Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA   Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA   Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT   "Comparative genomics of Listeria species.";
RL   Science 294:849-852(2001).
DR   EMBL; AL591982; CAD00226.1; -.
DR   ListiList; LMO02148; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   213 AA;  25361 MW;  A9EF9363E2FF0E78 CRC64;

  Query Match             80.4%;  Score 41;  DB 16;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           |||||||
Db      67 FPNAPYL 73


RESULT 10
Q98SJ4
ID   Q98SJ4      PRELIMINARY;      PRT;   250 AA.
```

```
AC   Q98SJ4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Wilms' tumor suppressor (Fragment).
GN   WT1.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Serluca F.C., Fishman M.C.;
RT   "Pre-pattern in the pronephric kidney field of zebrafish.";
RL   Development 0:0-0(2001).
DR   EMBL; AY028627; AAK27262.1; -.
DR   InterPro; IPR000976; Wilms_tumour.
DR   Pfam; PF02165; WT1; 1.
FT   NON_TER       1      1
FT   NON_TER     250    250
SQ   SEQUENCE   250 AA;  27236 MW;  90FA4B0B5FF59F6A CRC64;

  Query Match             78.4%;  Score 40;  DB 13;  Length 250;
  Best Local Similarity   77.8%;  Pred. No. 6.9;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
           ||| |:|||
Db      88 RMFSNSPYL 96


RESULT 11
Q8XPG7
ID   Q8XPG7      PRELIMINARY;      PRT;   289 AA.
AC   Q8XPG7;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical lipoprotein transmembrane protein RSp1673.
GN   RSP1673 OR RS02232.
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OG   Plasmid megaplasmid.
OC   Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC   Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
DR   EMBL; AL646086; CAD18824.1; -.
KW   Plasmid; Hypothetical protein; Complete proteome.
SQ   SEQUENCE   289 AA;  29665 MW;  7158677B3C03C9EC CRC64;

  Query Match             78.4%;  Score 40;  DB 16;  Length 289;
  Best Local Similarity   87.5%;  Pred. No. 7.9;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MFPNAPYL 9
           | ||||||
Db     115 MLPNAPYL 122


RESULT 12
Q91030
```

```
ID   Q91030      PRELIMINARY;      PRT;   392 AA.
AC   Q91030;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Chick Wilm's tumour protein (Fragment).
GN   CWT1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=EMBRYO;
RX   MEDLINE=96068905; PubMed=7478606;
RA   Kent J., Coriat A.M., Sharpe P.T., Hastie N., van Heyningen V.;
RT   "The evolution of WT1 sequence and expression in the vertebrates.";
RL   Oncogene 11:1781-1792(1995).
DR   EMBL; X85731; CAA59736.1; -.
DR   HSSP; P08046; 1A1G.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF02165; WT1; 1.
DR   Pfam; PF00096; zf-C2H2; 3.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 3.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 3.
KW   DNA-binding; Metal-binding; Zinc-finger.
FT   NON_TER     392    392
SQ   SEQUENCE   392 AA;  43869 MW;  761F01D350E4EDBF CRC64;

  Query Match             78.4%;  Score 40;  DB 13;  Length 392;
  Best Local Similarity   77.8%;  Pred. No. 11;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          ||||  |||
Db    109 RMFPQRPYL 117


RESULT 13
Q9I8A1
ID   Q9I8A1      PRELIMINARY;      PRT;   414 AA.
AC   Q9I8A1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   WT1(-KTS) protein.
GN   WT1(-KTS).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RA   Kudo T.;
RT   "Chicken counterpart of Wilms' tumor suppressor gene 1.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB033633; BAA94793.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF02165; WT1; 1.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 4.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   414 AA;  46641 MW;  17EB22F8B428A2EF CRC64;

  Query Match             78.4%;  Score 40;  DB 13;  Length 414;
  Best Local Similarity   77.8%;  Pred. No. 11;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          ||||  |||
Db    109 RMFPQRPYL 117


RESULT 14
Q9I8A0
ID   Q9I8A0      PRELIMINARY;      PRT;   417 AA.
AC   Q9I8A0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   WT1(+KTS) protein.
GN   WT1(+KTS).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RA   Kudo T.;
RT   "Chicken counterpart of Wilms' tumor suppressor gene 1.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB033634; BAA94794.1; -.
DR   HSSP; P08046; 1AAY.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF02165; WT1; 1.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 4.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   417 AA;  46957 MW;  47BBF7F6448E7F6C CRC64;

  Query Match             78.4%;  Score 40;  DB 13;  Length 417;
  Best Local Similarity   77.8%;  Pred. No. 11;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          ||||  |||
Db    109 RMFPQRPYL 117


RESULT 15
Q9PUT7
ID   Q9PUT7      PRELIMINARY;      PRT;   419 AA.
AC   Q9PUT7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Wilms' tumor suppressor.
GN   WT1.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith S.I., Down M., Power M., Boyd A.W.;
```

```
RT   "Isolation and characterization of a cDNA encoding zebrafish (Danio
RT   rerio) WT-1.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF144550; AAF00123.1; -.
DR   HSSP; P08046; 1AAY.
DR   ZFIN; ZDB-GENE-980526-558; wt1.
DR   InterPro; IPR000976; Wilms_tumour.
DR   InterPro; IPR000822; Znf_C2H2.
DR   Pfam; PF02165; WT1; 1.
DR   Pfam; PF00096; zf-C2H2; 4.
DR   PRINTS; PR00049; WILMSTUMOUR.
DR   ProDom; PD000003; Znf_C2H2; 2.
DR   SMART; SM00355; ZnF_C2H2; 4.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 4.
KW   DNA-binding; Metal-binding; Zinc-finger.
SQ   SEQUENCE   419 AA;  46925 MW;  606ADFEDA619EECD CRC64;

  Query Match             78.4%;  Score 40;  DB 13;  Length 419;
  Best Local Similarity   77.8%;  Pred. No. 12;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 RMFPNAPYL 9
          ||| |:|||
Db    111 RMFSNSPYL 119


Search completed: April 25, 2003, 06:48:34
Job time : 32 secs
```

```
                    GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         April 25, 2003, 06:46:39 ; Search time 44 Seconds
                                           (without alignments)
                                           19.664 Million cell updates/sec

Title:          US-09-625-963-1
Perfect score:  51
Sequence:       1 RMFPNAPYL 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters:      283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_73:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result           Query
  No.    Score   Match Length DB  ID                Description
-----------------------------------------------------------------------------
    1      51   100.0    410  2  JC5046             Wilms' tumor suppr
    2      51   100.0    448  2  S33926             Wilms' tumor prote
    3      51   100.0    449  2  A38080             Wilms tumor suscep
    4      51   100.0    449  2  A39692             Wilms' tumor prote
    5      41    80.4    213  2  AD1343             hypothetical prote
    6      41    80.4    213  2  AH1713             hypothetical prote
    7      40    78.4    262  1  G64144             hypothetical prote
    8      39    76.5    203  2  C71375             conserved hypothet
    9      38    74.5    205  2  F83495             probable hydrolase
   10      38    74.5    214  2  AB2195             serine esterase [i
   11      38    74.5    581  2  E86408             F3H9.11 protein -
   12      37    72.5    466  2  S42210             ISG-54K protein -
   13      37    72.5    854  2  S33558             unc-33 protein - C.
   14      37    72.5   1232  2  A55478             neuronal apoptosis
   15      36    70.6     85  2  S17463             glutathione transf
   16      36    70.6    115  2  S39393             glutathione transf
   17      36    70.6    142  2  S17462             glutathione transf
   18      36    70.6    217  2  JX0095             glutathione transf
   19      36    70.6    218  1  XURTG4             glutathione transf
   20      36    70.6    218  2  B34159             glutathione transf
   21      36    70.6    218  2  A29036             glutathione transf
   22      36    70.6    218  2  S13202             glutathione transf
   23      36    70.6    218  2  S33860             glutathione transf
   24      36    70.6    218  2  A46143             mu-class glutathio
   25      36    70.6    218  2  A23732             glutathione transf
   26      36    70.6    218  2  A39375             glutathione transf
   27      36    70.6    218  2  S32425             glutathione transf
   28      36    70.6    218  2  A46048             glutathione transf
   29      36    70.6    218  2  S01719             glutathione transf
   30      36    70.6    218  2  B28946             glutathione transf
   31      36    70.6    218  2  A29794             glutathione transf
   32      36    70.6    218  2  A47486             glutathione transf
   33      36    70.6    218  2  S65674             glutathione transf
   34      36    70.6    220  2  S18464             glutathione transf
   35      36    70.6    225  2  A35295             glutathione transf
   36      36    70.6    284  2  F95120             hydrolase, probabl
   37      36    70.6    284  2  B97990             conserved hypothet
   38      36    70.6    358  2  JC5964             apoptosis inhibito
   39      36    70.6    381  2  S65212             hypothetical prote
   40      36    70.6    382  2  A48492             polysaccharide exp
   41      36    70.6    432  2  B96515             hypothetical prote
   42      36    70.6    434  2  C96515             hypothetical prote
   43      36    70.6    487  2  T07960             probable (S)-N-met
   44      36    70.6    488  2  T07963             probable (S)-N-met
   45      36    70.6    618  2  S68450             apoptosis inhibito

                              ALIGNMENTS


RESULT 1
JC5046
Wilms' tumor suppressor protein - African clawed frog
N;Alternate names: WT1
C;Species: Xenopus laevis (African clawed frog)
C;Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 31-Jan-1997
C;Accession: JC5046
R;Semba, K.; Saito-Ueno, R.; Takayama, G.; Kondo, M.
Gene 175, 167-172, 1996
A;Title: cDNA cloning and its pronephros-specific expression of the Wilms' tumor supp
A;Reference number: JC5046; MUID:97074667; PMID:8917094
A;Contents: testis
A;Accession: JC5046
A;Molecule type: mRNA
A;Residues: 1-410 <SEM>
A;Cross-references: DDBJ:D82051
C;Comment: This protein is involved in kidney morphogenesis.
C;Genetics:
A;Gene: wt1

  Query Match             100.0%;  Score 51;  DB 2;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 0.049;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPYL 9
             |||||||||
Db     107 RMFPNAPYL 115


RESULT 2
S33926
Wilms' tumor protein WT1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 10-May-1996
C;Accession: S33926
R;Sharma, P.M.; Yang, X.; Bowman, M.; Roberts, V.; Sukumar, S.
Cancer Res. 52, 6407-6412, 1992
A;Title: Molecular cloning of rat Wilms' tumor complementary DNA and a study of messe
A;Reference number: S33926; MUID:93046155; PMID:1330293
A;Accession: S33926
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-448 <SHA>
A;Cross-references: EMBL:X69716
C;Genetics:
A;Gene: WT1
C;Keywords: tumor suppressor

  Query Match             100.0%;  Score 51;  DB 2;  Length 448;
  Best Local Similarity   100.0%;  Pred. No. 0.054;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        1 RMFPNAPYL 9
            |||||||||
Db      125 RMFPNAPYL 133


RESULT 3
A38080
Wilms tumor susceptibility protein WT1 - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Jun-2000
C;Accession: A38080; S08273; A34673; I38504; I52811; I58315; A56411; S26286
R;Gessler, M.; Konig, A.; Bruns, G.A.
Genomics 12, 807-813, 1992
A;Title: The genomic organization and expression of the WT1 gene.
A;Reference number: A38080; MUID:92241883; PMID:1572653
A;Accession: A38080
A;Molecule type: DNA
A;Residues: 1-449 <GES1>
A;Cross-references: GB:X61631; GB:S99414; NID:g37981; PIDN:CAA43819.1; PID:g825731
A;Note: it is uncertain whether Met-1 is the initiator or whether translation is initiat
A;Note: sequence extracted from NCBI backbone (NCBIN:99414, NCBIN:99422, NCBIN:99479, NC
A;Note: the sequence in GenBank entry HSWTGEEX1, release 113.0, PIDN:CAA43819.1 differs
R;Gessler, M.; Poustka, A.; Cavenee, W.; Neve, R.L.; Orkin, S.H.; Bruns, G.A.P.
Nature 343, 774-778, 1990
A;Title: Homozygous deletion in Wilms tumours of a zinc-finger gene identified by chrom
A;Reference number: S08273; MUID:90158822; PMID:2154702
A;Accession: S08273
A;Molecule type: mRNA
A;Residues: 'SRQRPHPGALRNPTACPLPHFPPSLPPTHSPTHPPRAGTAAQAPGPRRLLAAILDFLLLQDPASTCVPEPASQHT
A;Cross-references: EMBL:X51630; NID:g37977; PIDN:CAA35956.1; PID:g37978
R;Call, K.M.; Glaser, T.; Ito, C.Y.; Buckler, A.J.; Pelletier, J.; Haber, D.A.; Rose, E.
Cell 60, 509-520, 1990
A;Title: Isolation and characterization of a zinc finger polypeptide gene at the human
A;Reference number: A34673; MUID:90150277; PMID:2154335
A;Accession: A34673
A;Molecule type: mRNA
A;Residues: 85-249,267-364,'F',366-386,'T',388-407,411-449 <CAL>
A;Cross-references: GB:M30393; NID:g340381; PIDN:AAA36810.1; PID:g340382
R;Haber, D.A.; Sohn, R.L.; Buckler, A.J.; Pelletier, J.; Call, K.M.; Housman, D.E.
Proc. Natl. Acad. Sci. U.S.A. 88, 9618-9622, 1991
A;Title: Alternative splicing and genomic structure of the Wilms tumor gene WT1.
A;Reference number: A56411; MUID:92052142; PMID:1658787
A;Contents: annotation; alternative splicing
R;Phelan, S.A.; Lindberg, C.; Call, K.M.
Cell Growth Differ. 5, 677-686, 1994
A;Title: Wilms' tumor gene, WT1, mRNA is down-regulated during induction of erythroid a
A;Reference number: I38504; MUID:94368704; PMID:8086342
A;Accession: I38504
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-18 <PHE>
A;Cross-references: EMBL:U06486; NID:g473563; PIDN:AAA62865.1; PID:g458432
R;Pelletier, J.; Bruening, W.; Kashtan, C.E.; Mauer, S.M.; Manivel, J.C.; Striegel, J.E.
Cell 67, 437-447, 1991
A;Title: Germline mutations in the Wilms' tumor suppressor gene are associated with abn
A;Reference number: I52811; MUID:92005721; PMID:1655284
A;Accession: I52811
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 355-365,'H',367-377 <PEL>
A;Cross-references: GB:S61513; NID:g237599; PIDN:AAB20109.1; PID:g237600
A;Note: mutant form
R;Hamilton, T.B.; Barilla, K.C.; Romaniuk, P.J.
Nucleic Acids Res. 23, 277-284, 1995
A;Title: High affinity binding sites for the Wilms' tumour suppressor protein WT1.
A;Reference number: I58315; MUID:95166649; PMID:7862533
A;Accession: I58315
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 'MGHHHHHHHHHHSSGHIEGRHM',301-364,'F',366-386,'T',388-407,411-449 <HAM>
A;Cross-references: GB:S75264; NID:g896246; PIDN:AAB33443.1; PID:g896247
A;Note: this sequence is engineered
C;Genetics:
```

```
A;Gene: GDB:WT1
A;Cross-references: GDB:120496; OMIM:194070
A;Map position: 11p13-11p13
A;Introns: 148/1; 189/1; 223/2; 249/2; 266/2; 298/3; 349/1; 379/1; 410/1
A;Note: mRNA transcripts containing both alternatively spliced regions are the most a
C;Keywords: alternative splicing; DNA binding; kidney; tumor suppressor; zinc finger
F;1-449/Product: Wilms tumor susceptibility protein WT1, splice form 1 #status predic
F;1-407,411-449/Product: Wilms tumor susceptibility protein WT1, splice form 4 #statu
F;1-249,267-449/Product: Wilms tumor susceptibility protein WT1, splice form 3 #statu
F;1-249,267-407,411-449/Product: Wilms tumor susceptibility protein WT1, splice form

  Query Match             100.0%;  Score 51;  DB 2;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.054;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db      126 RMFPNAPYL 134


RESULT 4
A39692
Wilms' tumor protein analog, WT1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 16-Feb-1997
C;Accession: A39692
R;Buckler, A.J.; Pelletier, J.; Haber, D.A.; Glaser, T.; Housman, D.E.
Mol. Cell. Biol. 11, 1707-1712, 1991
A;Title: Isolation, characterization, and expression of the murine Wilms' tumor gene
A;Reference number: A39692; MUID:91141522; PMID:1671709
A;Accession: A39692
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-449 <BUC>
A;Cross-references: GB:M55512
C;Keywords: alternative splicing; DNA binding; transcription regulation; tumor suppre

  Query Match             100.0%;  Score 51;  DB 2;  Length 449;
  Best Local Similarity   100.0%;  Pred. No. 0.054;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 RMFPNAPYL 9
            |||||||||
Db      126 RMFPNAPYL 134


RESULT 5
AD1343
hypothetical protein lmo2148 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AD1343
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AD1343
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-213 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00226.1; PID:g16411618; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2148

  Query Match              80.4%;  Score 41;  DB 2;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy       3 FPNAPYL 9
           |||||||
Db      67 FPNAPYL 73


RESULT 6
AH1713
hypothetical protein homolog lin2252 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AH1713
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1713
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-213 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97480.1; PID:g16414764; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2252

  Query Match             80.4%;  Score 41;  DB 2;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 1.9;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           |||||||
Db      67 FPNAPYL 73


RESULT 7
G64144
hypothetical protein HI0177 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jul-2000
C;Accession: G64144
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800
A;Accession: G64144
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-262 <TIGR>
A;Cross-references: GB:U32703; GB:L42023; NID:g1573133; PIDN:AAC21847.1; PID:g1573134; T
A;Note: best homolog was a hypothetical protein from Pseudomonas aeruginosa
C;Superfamily: conserved hypothetical protein HI0177

  Query Match             78.4%;  Score 40;  DB 1;  Length 262;
  Best Local Similarity   75.0%;  Pred. No. 3.8;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 RMFPNAPY 8
           |:|||:||
Db     151 RVFPNSPY 158


RESULT 8
C71375
conserved hypothetical integral membrane protein TP0033 - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C;Accession: C71375
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: C71375
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-203 <COL>
A;Cross-references: GB:AE001188; GB:AE000520; NID:g3322282; PIDN:AAC65028.1; PID:g332
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0033

  Query Match             76.5%;  Score 39;  DB 2;  Length 203;
  Best Local Similarity   85.7%;  Pred. No. 4.4;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
           ||||||:
Db      63 FPNAPYI 69


RESULT 9
F83495
probable hydrolase PA1202 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: F83495
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83495
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-205 <STO>
A;Cross-references: GB:AE004550; GB:AE004091; NID:g9947122; PIDN:AAG04591.1; GSPDB:GN
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1202

  Query Match             74.5%;  Score 38;  DB 2;  Length 205;
  Best Local Similarity   75.0%;  Pred. No. 7;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 MFPNAPYL 9
            ||| |||:
Db      75 MFPQAPYI 82


RESULT 10
AB2195
serine esterase [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp.
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C;Accession: AB2195
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AB2195
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-214 <KUR>
A;Cross-references: GB:BA000019; PIDN:BAB74812.1; PID:g17132207; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
```

```
A;Gene: all3113

  Query Match             74.5%;  Score 38;  DB 2;  Length 214;
  Best Local Similarity   85.7%;  Pred. No. 7.3;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 MFPNAPY 8
              :||||||
Db         56 VFPNAPY 62


RESULT 11
E86408
F3H9.11 protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Nov-2001
C;Accession: E86408
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E86408
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-581 <STO>
A;Cross-references: GB:AE005172; NID:g9795613; PIDN:AAF98431.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

  Query Match             74.5%;  Score 38;  DB 2;  Length 581;
  Best Local Similarity   87.5%;  Pred. No. 22;
  Matches    7;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPY 8
              |||||| |
Db        349 RMFPNARY 356


RESULT 12
S42210
ISG-54K protein - Chinese hamster
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 07-May-1999
C;Accession: S42210
R;Bluyssen, H.A.R.; Vlietstra, R.J.; van der Made, A.; Trapman, J.
Eur. J. Biochem. 220, 395-402, 1994
A;Title: The interferon-stimulated gene 54 K promoter contains two adjacent functional i
alpha inducibility.
A;Reference number: S42210; MUID:94170787; PMID:8125096
A;Accession: S42210
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-466 <BLU>
C;Superfamily: interferon-induced 56K protein

  Query Match             72.5%;  Score 37;  DB 2;  Length 466;
  Best Local Similarity   66.7%;  Pred. No. 27;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              || ||:||:
Db        201 RMSPNSPYV 209


RESULT 13
S33558
```

```
unc-33 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 07-Apr-1994 #sequence_revision 02-Jun-1994 #text_change 21-Jul-2000
C;Accession: S33558; S33559; S24643; S24644
R;Li, W.; Herman, R.K.; Shaw, J.E.
Genetics 132, 675-689, 1992
A;Title: Analysis of the Caenorhabditis elegans axonal guidance and outgrowth gene un
A;Reference number: S33558; MUID:93106371; PMID:1468626
A;Accession: S33558
A;Molecule type: DNA
A;Residues: 1-854 <LIW1>
A;Cross-references: EMBL:Z14148; NID:g6899; PIDN:CAA78520.1; PID:g6900
A;Accession: S33559
A;Molecule type: mRNA
A;Residues: 1-854 <LIW2>
A;Cross-references: EMBL:Z14146; NID:g6903; PID:g6904
C;Genetics:
A;Introns: 14/2; 68/1; 129/3; 151/3; 182/1; 255/3; 361/2; 660/1; 790/3
C;Keywords: alternative initiators
F;1-854/Product: unc-33 protein (long form) #status predicted <MAT1>
F;176-854/Product: unc-33 protein (intermediate form) #status predicted <MAT2>
F;332-854/Product: unc-33 protein (short form) #status predicted <MAT3>

  Query Match             72.5%;  Score 37;  DB 2;  Length 854;
  Best Local Similarity   66.7%;  Pred. No. 52;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              |: ||:|||
Db        773 RLSPNSPYL 781


RESULT 14
A55478
neuronal apoptosis inhibitory protein - human
N;Alternate names: NAIP
C;Species: Homo sapiens (man)
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 02-Feb-2001
C;Accession: A55478
R;Roy, N.; Mahadevan, M.S.; McLean, M.; Shutler, G.; Yaraghi, Z.; Farahani, R.; Baird
d, T.O.; de Jong, P.J.; Surh, L.; Ikeda, J.E.; Korneluk, R.G.; MacKenzie, A.
Cell 80, 167-178, 1995
A;Title: The gene for neuronal apoptosis inhibitory protein is partially deleted in i
A;Reference number: A55478; MUID:95112344; PMID:7813013
A;Accession: A55478
A;Molecule type: mRNA
A;Residues: 1-1232 <ROY>
A;Cross-references: GB:U19251
C;Genetics:
A;Gene: GDB:SMA@; SMA
A;Cross-references: GDB:120378; OMIM:600354; OMIM:253300
A;Map position: 5q12.2-5q13
C;Keywords: apoptosis; ATP; glycoprotein; nucleotide binding; P-loop; transmembrane p
F;94-110/Domain: transmembrane #status predicted <TMM1>
F;470-477/Region: nucleotide-binding motif A (P-loop)
F;479-496/Domain: transmembrane #status predicted <TMM2>
F;476/Binding site: ATP (Lys) #status predicted
F;618,632,823,923,1035/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             72.5%;  Score 37;  DB 2;  Length 1232;
  Best Local Similarity   66.7%;  Pred. No. 79;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 RMFPNAPYL 9
              | ||| |:|
Db        337 RCFPNCPFL 345


RESULT 15
S17463
glutathione transferase (EC 2.5.1.18) class mu chain Yb5 - mouse (fragments)
C;Species: Mus musculus (house mouse)
C;Date: 22-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 16-Feb-1997
```

```
C;Accession: S17463
R;Hayes, J.D.; Kerr, L.A.; Peacock, S.D.; Cronshaw, A.D.; McLellan, L.I.
Biochem. J. 277, 501-512, 1991
A;Title: Hepatic glutathione S-transferases in mice fed on a diet containing the antical
adient elution of the glutathione-Sepharose affinity matrix.
A;Reference number: S16933; MUID:91315425; PMID:1859377
A;Accession: S17463
A;Molecule type: protein
A;Residues: 1-38;39-64;65-85 <HAY>
C;Superfamily: glutathione transferase
C;Keywords: blocked amino end; dimer; transferase

  Query Match            70.6%;  Score 36;  DB 2;  Length 85;
  Best Local Similarity  85.7%;  Pred. No. 6.3;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 FPNAPYL 9
             ||| |||
Db      23 FPNLPYL 29


Search completed: April 25, 2003, 06:52:27
Job time : 46 secs
```